目 录

插图目录

表格目录

第1章 组播概述

📖 **说明：**

在组播业务中，MA5200G 作为组播路由器使用，下文中直接以组播路由器代指 MA5200G。

本章对组播的相关知识进行简单介绍，主要包括以下内容：

- 组播简介
- 组播地址
- 组播协议
- 组播转发
- 组播应用

1.1 组播简介

1.1.1 单播、广播和组播

单播、广播和组播是指信息源（Source）向接收者（Receiver）进行信息传送的方式。

1. 单播 Unicast

在单播方式下，信息源和每一个接收者之间都建立一条数据传送通路，并在独立的通路上进行信息传送。

单播方式适合用户稀少的网络。当信息源需要向大量接收者同时进行信息传送（例如在网络会议、视频点播、远程教学等应用）的时候，单播方式将占用大量的带宽资源。

2. 广播 Broadcast

在广播方式下，信息源不和接收者建立单独的数据传送通路，而是将信息在网络中进行广播，网络中所有的用户都能接收到。

广播方式适合用户稠密的网络。广播方式同样需要占用大量的带宽资源，另外广播方式也不利于信息的安全和保密。

3. 组播 Multicast

组播方式下，信息源仅发送一次信息，组播路由协议为组播数据报文建立树型路由，被传递的信息在尽可能远的分叉路口才开始复制和分发（参见图 1-1），因此，信息能够被准确高效地传送到每个需要它的用户。

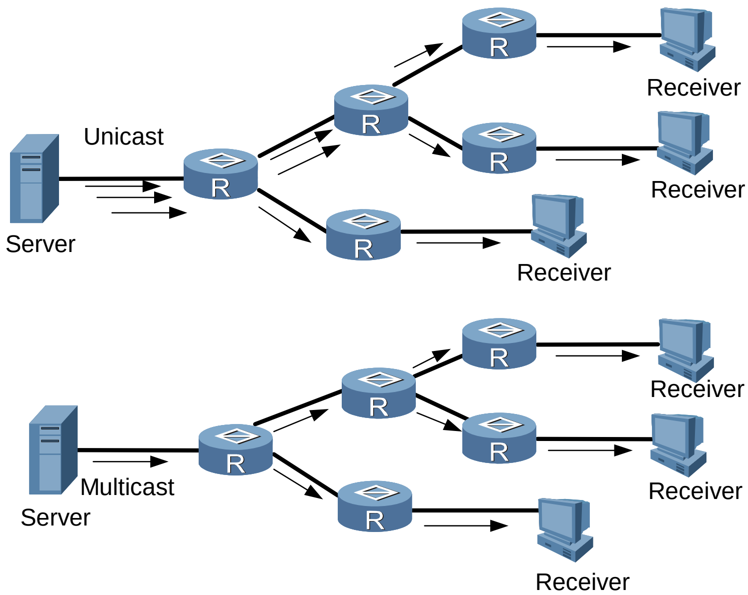


图1-1 单播与组播传送信息的对比

组播方式的出现，解决了信息传送中的带宽和安全问题。组播已经成为 IP 网络中信息传送的主流方式。相比单播和广播，组播主要有以下三个优势。

- 提高效率：降低网络流量，减轻服务器和 CPU 负荷；
- 优化性能：减少冗余流量；
- 分布式应用：使多点应用成为可能。

1.1.2 组播相关的术语

1. 组播源

组播方式中，信息源被称为组播源。

2. 组播组

组播方式中，接收相同信息的接收者构成一个组播组，并且每个接收者都是该组播组的成员。同一组播组的成员可以广泛分布在网络中的任何地方，即组播组关系没有地域限制。

组播组以组播地址进行标识，请参见“1.2 组播地址”。

📖 说明：

组播源不一定属于组播组，它向组播组发送数据，自己不一定是接收者。可以同时有多个源向一个组播组发送报文。

3. 组播路由器

进行组播协议处理、组播报文的复制和分发的路由器被称为组播路由器。

网络中可能有不支持组播的路由器，组播路由器可以使用隧道方式将组播报文封装在单播 IP 报文中传送给相邻的组播路由器，相邻的组播路由器再将单播 IP 头剥离，然后继续进行组播传输。

1.2 组播地址

1.2.1 IP 组播地址

组播报文的目的地址使用 D 类 IP 地址，范围是 224.0.0.0 到 239.255.255.255。D 类地址不能出现在 IP 报文的源 IP 地址字段。

单播数据传输过程中，一个数据报文传输的路径是从源地址路由到目的地址，利用“逐跳”（hop-by-hop）的原理在 IP 网络中传输。然而在 IP 组播环境中，数据报文的目的地址不是一个，而是一组，形成组地址。

所有的信息接收者都加入到一个组内，并且一旦加入之后，流向组地址的数据立即开始向接收者传输，组中的所有成员都能接收到数据报文。组播组中的成员是动态的，主机可以在任何时刻加入和离开组播组。

组播组可以是永久的也可以是临时的。组播组地址中，有一部分由官方分配的，称为永久组播组。永久组播组保持不变的是它的 IP 地址，组中的成员构成可以发生变化。永久组播组中成员的数量都可以是任意的，甚至可以为零。未保留的 IP 组播地址，可以被临时组播组利用。

组播组地址的范围及含义见表 1-1。

表1-1 组播组地址的范围及含义

D 类地址范围	含义
224.0.0.0~224.0.0.255	永久组播组地址。其中地址 224.0.0.0 保留不做分配，其他地址供路由协议使用。
224.0.1.0~238.255.255.255	临时组播组地址，全网范围内有效。
239.0.0.0~239.255.255.255	本地管理组播地址，仅在特定的本地范围内有效。

1.2.2 以太网组播 MAC 地址

以太网传输单播 IP 报文的时候，目的 MAC 地址使用的是接收者的 MAC 地址。但是在传输组播报文时，传输目的不再是一个具体的接收者，而是一个成员不确定的组，所以使用的是组播 MAC 地址。

组播 MAC 地址是和组播 IP 地址对应的。IANA 规定，组播 MAC 地址的高 24bit 为 0x01005e，第 25bit 为 0，MAC 地址的低 23bit 为组播 IP 地址的低 23bit。

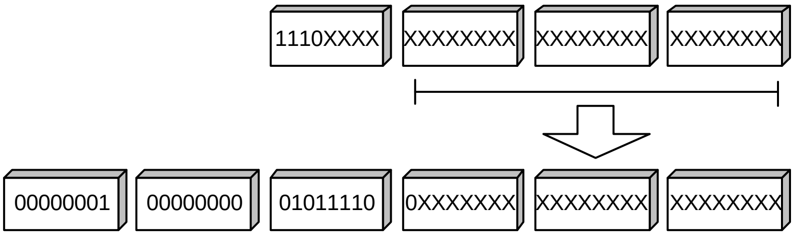


图1-2 组播 IP 地址与以太网 MAC 地址的映射关系

由于 IP 组播地址的后 28 位中只有低 23 位被映射到 MAC 地址，这样就会有 32 个 IP 组播地址映射到同一 MAC 地址上。

1.3 组播协议

组播涉及到的协议主要有组播组管理协议和组播路由协议。

组播组管理协议目前使用的是作为 IP 组播基本信令协议的因特网组管理协议 IGMP（Internet Group Management Protocol），它运行在主机和路由器之间，作用是使路由器了解网段上有没有组播组成员。

组播路由协议运行在组播路由器之间，作用是建立和维护组播路由，正确、高效地转发组播数据报文。常用的组播路由协议包括 PIM-DM（Protocol -Independent Multicast Dense Mode）、PIM-SM（Protocol-Independent Multicast Sparse Mode）、MSDP（Multicast Source Discovery Protocol）等，而 BGP 经过组播扩展后（MBGP）也可以在域间传递组播路由信息。

1.3.1 IGMP

IGMP 是主机可以使用的唯一协议，它定义了主机与路由器之间组播成员关系的建立和维护机制，是整个 IP 组播的基础。IGMP 的基本原理如下：

- 主机通过 IGMP 通知路由器有关组成员的信息，并使路由器能通过和自己直接相连的主机来了解组内其他成员的情况。

- 如果一个网络中有一个用户通过 IGMP 宣布加入某组播组，则网络中的组播路由器就将发到该组播组的信息通过组播路由协议进行传播，最终将该网络作为一个分枝加入组播树。
- 当主机作为某个组的成员开始收到信息后，路由器就会周期性地对该组进行查询，检查组内的成员是否还参与其中。只要还有一个主机仍在参与，路由器就继续接收数据。
- 当网络中的所有用户退出该组播组后，相关的分枝就从组播树中删掉。

1.3.2 组播路由协议

组播路由建立了一个从数据源端到多个接收端的无环数据传输路径。组播路由协议的任务就是构建分发树结构。组播路由器能采用多种方法来建立数据传输的路径，即分发树。

与单播路由一样，组播路由也分为域内和域间两大类。

- 域内组播路由协议：常用的包括 PIM-DM 和 PIM-SM。
- 域间组播路由协议：常用的包括 MSDP 和 MBGP。

1. PIM-DM

PIM-DM 协议适用于接收者相对密集的小型网络。其基本原理如下：

- PIM-DM 首先假设网络中的每个子网都存在至少一个对组播源感兴趣的接收站点。因此，组播数据报文被扩散到网络中的所有点。与此伴随着相关资源（带宽和路由器的 CPU 等）的消耗。
- 为了减少这些宝贵网络资源的消耗，PIM-DM 对没有组播数据转发的分支进行剪枝操作，只保留包含接收站点的分支。
- 为了使剪掉的分支中有组播数据转发需求的接收站点可以接收组播数据流，剪掉的分支可以周期性地恢复成转发状态。
- 为了减少等待剪枝分支恢复转发状态的延时时间，密集模式组播路由协议使用嫁接机制主动恢复组播报文的转发。

这种周期性的扩散和剪枝现象是密集模式协议的特征。一般说来，密集模式下数据报文的转发路径是有源树（以组播源为根、组播组成员为枝叶的一棵树）。由于有源树使用的是从组播源到接收者的最短路径，因此也称为最短路径树 SPT（Shortest Path Tree）。

2. PIM-SM

PIM-SM 适用于组播接收成员相对稀疏的广域网。其基本原理如下：

- PIM-SM 首先假设所有主机都不需要接收组播报文，只向明确指定需要组播报文的主机转发。
- 为了使接收者能够接收到特定组的组播数据流，连接这些接收者的组播路由器必须向该组对应的汇聚点 RP（Rendezvous Point）（汇聚点需要在网络中构建，是一些虚拟的数据交换地点）发送加入消息，加入消息经过一个个组播路由器后到达根部，即汇聚点，所经过的路径就变成了共享树的分支。
- 为了避免共享树的分支由于不被更新而被删除，稀疏模式组播路由协议通过向分支周期性地发送加入消息来维护组播分布树。
- 组播源首先要在汇聚点进行注册，之后把数据发向汇聚点。
- 当数据到达了汇聚点后，组播数据报文被复制并沿着分发树路径把数据传给接收者。复制仅仅发生在分发树的分支处，这个过程能自动重复直到数据报文最终到达目的地。

3. MSDP

对于 ISP 来说，不希望依靠竞争对手的 RP 转发组播流量，但同时又要求无论组播源注册的 RP 在哪里，都能从组播源获取信息发给自己内部的成员。

MSDP 就是为了解决这个问题而提出的。MSDP 用于多个 PIM-SM 域互连。MSDP 允许不同域的 RP 共享其组播源信息。MSDP 要求域内组播路由协议必须是 PIM-SM。

4. MBGP 组播扩展

当前使用最多的域间单播路由协议是 BGP-4。由于组播的网络拓扑和单播拓扑有可能不同，为了实现域间组播路由信息的传递，必须对 BGP-4 进行改造。为了构造域间组播路由树，除了要知道单播路由信息外，还要知道网络中哪些部分是支持组播的，即组播的网络拓扑情况。

在 RFC2858 中规定了对 BGP 进行多协议扩展的方法，扩展后的 BGP（MBGP，也写作 BGP-4+）不仅能携带 IPv4 单播路由信息，也能携带其他网络层协议（如组播、IPv6 等）的路由信息，携带组播路由信息只是其中一个扩展功能。

1.4 组播转发

在组播方式中，组播源向组播组传送信息。和单播不同，组播路由器必须将组播数据报文转发到多个外部接口上以便能传送到所有接收站点，因此组播转发过程比单播转发过程更加复杂。

1.4.1 逆向路径转发 RPF

为了保证组播报文都是通过最短路径到达路由器，组播必须依靠单播路由表或者单独提供给组播使用的单播路由表（如 MBGP 组播路由表），对组播报文的接收接口进行 RPF（Reverse Path Forwarding）检查。RPF 检查是大部分组播路由协议进行组播转发的基础。

组播路由器以组播报文源地址来查询单播路由表或者独立的组播路由表，以确定此数据报文到达的入接口处于接收站点至源地址的最短路径上。

- 如果使用的是有源树，这个源地址就是发送组播数据包的组播源的地址；
- 如果使用的是共享树，该源地址就是共享树的根的地址，即 RP。

当组播数据包到达路由器时，如果 RPF 检查通过，数据包则按照组播转发项进行转发，否则数据包被丢弃。

1.4.2 组播策略路由

组播策略路由是对组播通常按照路由表进行报文转发功能的一种补充和增强，它依照用户指定的具体策略来转发组播报文。

组播策略路由通过配置 Route-policy 来实现，它是单播策略路由的一种扩展，由用户输入的一组 If-match-apply 语句来描述。If-match 子句定义匹配准则，也就是通过当前 Route-policy 规定所需满足的过滤条件，它规定当组播报文满足用户设置的匹配条件时，不再按照通常的流程来转发，而是按照用户设置的动作（由 apply 语句描述）进行转发。

1.5 组播复制

对于广播方式的 BTV（Broadcast TV）业务，MA5200G 可以提供面向用户的组播复制功能，包括面向用户的 PVC/VLAN 或者 PPP Session 进行 BTV 业务的组播复制。运营商可以通过网管或者 AAA 对用户的 BTV 业务进行授权控制，保证 IPTV 业务的可运营性。

1.6 组播应用

组播方式有效地解决了单点发送多点接收的问题，实现了 IP 网络中点到多点的高效数据传送，能够大量节约网络带宽、降低网络负载。利用组播方式可以方便地提供一些新的增值业务，简单描述如下：

- 多媒体、流媒体的应用；
- 培训、联合作业场合的通信；

- 数据仓库、金融应用（股票）等；
- 任何“点到多点”的数据发布应用。

在 IP 网络中多媒体业务日渐增多的情况下，组播有着巨大的市场潜力，组播业务也将逐渐得到推广和普及。

第2章 配置组播公共信息

本章介绍组播公共信息的配置方法，这些配置与具体的组播协议无关。本章包括以下内容：

- 配置内容
- 显示和调试

2.1 配置内容

配置组播公共信息主要包括以下内容：

- 启动组播
- 配置组播转发报文的最小 TTL 值
- 配置组播报文的 VLAN/PVC
- 配置组播路由容量限制
- 配置组播复制功能

2.1.1 启动组播

启动组播路由协议之前，必须首先启动组播。全局启动组播是各接口启用组播功能的前提条件。

请在系统视图下进行下列配置。

操作	命令
启动组播	**multicast routing-enable**
停止组播	**undo multicast routing-enable**

缺省情况下，不启动组播。

注意：

只有启动了组播，其他的组播配置才能生效。

2.1.2 配置组播转发报文的最小 TTL 值

组播转发 TTL（Time to Live）值可配置在所有支持组播报文转发的接口上。

在接口上配置组播转发报文的最小 TTL 值后，当要将一个报文从该接口转发出去时（包括本机发出的报文），对接口上配置的最小 TTL 值进行检查，若报文 TTL 值（报文 TTL 已在本路由器内减 1）大于接口上配置的最小 TTL 值，则转发该报文；若报文 TTL 值小于或等于接口上配置的最小 TTL 值，则丢弃该报文。

请在接口视图下进行下列配置。

操作	命令
配置组播转发报文的最小 TTL 值	**multicast minimum-ttl** *ttl-value*
取消组播转发报文的最小 TTL 值	**undo multicast minimum-ttl**

缺省情况下，未配置组播转发的最小 TTL 值。

2.1.3 配置组播报文的 VLAN/PVC

配置组播报文的 VLAN/PVC 是为了把组播报文和其他报文进行隔离，保证组播报文的安全性。

组播报文的 VLAN/PVC 是基于接口进行配置。您可以在以太网接口或以太网 Trunk 接口上配置该接口的组播报文的 VLAN，而在 ATM 接口上配置该接口的组播报文的 PVC。

说明：

- 您只能在以太网、Trunk、ATM 接口的主接口上配置组播报文的 VLAN/PVC，而不能在这些接口的子接口上进行配置。
- 建议组播报文的 VLAN/PVC 不要和该接口下的用户报文的 VLAN/PVC 相同，CLI（Command Line Interface）对此未做限制，必须由用户在配置时予于保证。
- 组播报文的 PVC 必须是已经存在的 PVC。

请在接口视图下进行下列配置。

操作	命令
配置组播报文的 VLAN	**multicast-vlan** *vlan-id* [**qinq** *qinq-vlan*]
取消组播报文的 VLAN	**undo multicast-vlan**
配置组播报文的 PVC	**multicast-pvc** *vpi/vci*
取消组播报文的 PVC	**undo multicast-pvc**

缺省情况下，未配置组播报文的 VLAN/PVC。

说明：

要使用本功能，MA5200G 下挂的 LAN Switch 或 DSLAM 设备必须支持按 VLAN/PVC 复制组播报文的功能。

2.1.4 配置组播路由容量限制

通过限制组播路由的容量，MA5200G 可以节约内存等资源。当组播路由数达到容量限制后，MA5200G 将丢弃新增（S，G）的数据和协议报文。

请在系统视图下进行下列配置。

操作	命令
配置组播路由容量限制	**multicast route-limit** *limit*
恢复缺省值	**undo multicast route-limit**

缺省情况下，组播路由的容量限制为 8192。

如果在配置该命令时，路由表中的路由已经超过配置值，原路由表中的路由不会被删除，系统提示“当前的路由数量比配置的多”。

2.1.5 配置组播复制功能

通常情况下，MA5200G 收到某个组播组的组播报文后，只会向每个物理端口复制一份组播报文，二层设备再将组播报文复制给该组播组的每个用户。

如果二层设备不具备 IGMP Snooping 功能，无法识别组播组用户，则需要在 MA5200G 的接口上启用按会话进行组播复制的功能，由 MA5200G 直接将组播报文复制到每个用户的会话中，发送给用户。

请在接口视图下进行下列配置。

操作	命令
启用按会话进行组播复制功能	**multicast copy by-session**
关闭按会话进行组播复制功能	**undo multicast copy**

缺省情况下，接口未启用按会话进行组播复制的功能。

2.2 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后组播的运行情况，验证配置的效果。

在用户视图下，执行 **debugging** 命令可对组播进行调试。

操作	命令
查看组播路由表信息	**display multicast routing-table** [*group-address* [**mask** { *mask* | *mask-length* }] | *source-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** }]*
查看组播转发表信息	**display multicast forwarding-table** [*group-address* [**mask** { *mask* | *mask-length* }] | *source-address* [**mask** { *mask* | *mask-length* }] | **incoming-interface** { *interface-type interface-number* | **register** }]*
查看 RPF 路由信息	**display multicast rpf-info** *source-address*
打开组播报文转发调试信息开关	**debugging multicast forwarding**
关闭组播报文转发调试信息开关	**undo debugging multicast forwarding**
打开组播转发状态调试信息开关	**debugging multicast status-forwarding**
关闭组播转发状态调试信息开关	**undo debugging multicast status-forwarding**
打开组播核心路由调试信息开关	**debugging multicast kernel-routing**
关闭组播核心路由调试信息开关	**undo debugging multicast kernel-routing**

第3章 配置 IGMP

本章包括以下内容：

- IGMP 简介
- 配置内容
- 显示和调试
- 故障排除

3.1 IGMP 简介

3.1.1 IGMP 概述

IGMP 是 TCP/IP 协议簇中负责 IP 组播成员管理的协议。它用来在 IP 主机和与其直接相邻的组播路由器之间建立、维护组播组成员关系。图 3-1描述了 IGMP 在 IP 组播中的应用位置。

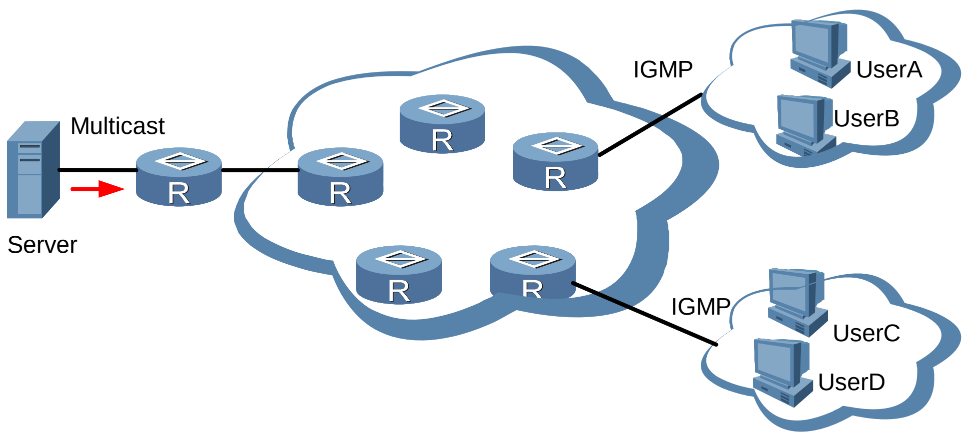


图3-1 IGMP 的应用位置

IGMP 有 IGMPv1（RFC1112）、IGMPv2（RFC2236）和 IGMPv3 三个版本。目前应用最多的是 v2 版本。

3.1.2 IGMPv1 的基本机制

IGMPv1 主要基于查询/响应机制完成组播组管理。IGMP 借助三层路由协议选举 DR 路由器，由 DR 路由器负责发送查询消息，主机加入组播组过程如下：

(1) DR 周期性地向共享网络内所有主机以组播方式发送成员组 Query 查询消息。

(2) 网络内所有主机都接收到该查询，属于该组播组的某主机以组播方式发送 Report 报告进行响应，从而加入该组播组。

(3) 网络中的所有主机和路由器都接收到该 Report 报告，如果某主机也属于该组播组则不再重复发送相关的 Report 报告；如果某主机属于另一组播组则发送 Report 报告进行响应，加入另一组播组。

(4) 经过 Query/Report 过程后，网络内支持组播的路由器就了解到本网络内哪些组播组包含成员，从而决定是否进行组播信息转发。

由于 IGMPv1 版本没有定义专门的离开组播组消息，因此当主机离开组播组时，主要基于查询无响应的思路实现。

3.1.3 IGMPv2 的基本机制

IGMPv2 版本对 v1 版本所做的改进主要有：

(1) 共享网段上组播路由器的选举机制

共享网段即一个网段上有多个组播路由器的情况。在这种情况下，由于此网段下运行 IGMP 的路由器都能从主机那里收到成员资格报告消息，因此，只需要一个路由器发送成员资格查询消息，这就需要一个路由器选举机制来确定一个路由器作为 DR。在 IGMPv1 版本中，DR 的选择由组播路由协议决定；IGMPv2 版本对此做了改进，规定同一网段上有多个组播路由器时，具有最低 IP 地址的组播路由器被选举出来充当 DR。

(2) 增加了离开组机制

在 IGMPv1 版本中，主机悄然离开组播组，不会给任何组播路由器发出任何通知。造成组播路由器只能依靠组播组响应超时来确定组播成员的离开。而在 v2 版本中，当一个主机决定离开一个组播组时，如果它是对最近一条成员资格查询消息作出响应的主机，那么它就会发送一条离开组的消息。

(3) 增加了对特定组的查询

在 IGMPv1 版本中，组播路由器的一次查询，是针对该网段下的所有组播组。这种查询称为普遍组查询。在 IGMPv2 版本中，在普遍组查询之外增加了特定组的查询，这种查询报文的目的 IP 地址为该组播组的 IP 地址，报文中的组地址域部分也为该组播组的 IP 地址。这样就避免了属于其他组播组成员的主机发送响应报文。

(4) 增加了最大响应时间字段

IGMPv2 版本增加最大响应时间字段，以动态地调整主机对组查询报文的响应时间。

3.2 配置内容

IGMP 的配置包括以下内容：

- 启动 IGMP
- 配置 IGMP 版本
- 配置 IGMP 查询报文间隔
- 配置 IGMP DR 存在时间
- 配置 IGMP 最大查询响应时间
- 配置路由器加入指定组播组
- 控制接口对组播组的访问
- 启用 IGMP 组快速离开功能

对于 IGMP 配置，只有启动 IGMP 是必选的，其他项都是可选的。

3.2.1 启动 IGMP

在需要进行组播成员关系维护的接口上使能 IGMP。执行此操作之后，才能进行 IGMP 的其他配置。

请在接口视图下进行下列配置。

操作	命令
启动 IGMP	**igmp enable**
停止 IGMP	**undo igmp enable**

3.2.2 配置 IGMP 版本

请在接口视图下进行下列配置。

操作	命令
配置路由器使用的 IGMP 版本	**igmp version** { **1** | **2** }
恢复缺省值	**undo igmp version**

缺省时，使用版本 2。

注意：

IGMP 各版本之间不能自动转换。因此，应该配置子网上的所有路由器使用同一 IGMP 版本。

3.2.3 配置 IGMP 查询报文间隔

路由器通过在所连接的网段上发送 IGMP 查询报文，获得该网段哪些组播组有成员。组播路由器周期性地发送查询报文，当得到某一组成员报告的时候，刷新该网段相应的组成员关系信息。

请在接口视图下进行下列配置。

操作	命令
配置 IGMP 查询报文间隔	**igmp timer query** *seconds*
恢复缺省值	**undo igmp timer query**

在一个网段中有多个组播路由器时，由查询器负责向局域网上的所有主机发送 IGMP 查询报文。

缺省情况下，IGMP 查询报文的时间间隔为 60 秒。

3.2.4 配置 IGMP DR 存在时间

IGMP DR 存在时间是指当 DR 停止发送查询报文后，另一路由器取代其成为新的 DR 的等待时间。

请在接口视图下进行下列配置。

操作	命令
配置 IGMP DR 存在时间	**igmp timer other-querier-present** *seconds*
恢复缺省值	**undo igmp timer other-querier-present**

缺省值为 120 秒。缺省时，路由器在 **igmp timer query** 命令指定的间隔的 2 倍时间内没有收到查询消息，就认为原来的 DR 失效。

3.2.5 配置 IGMP 最大查询响应时间

从路由器接收到一条查询消息后，主机会为其所属的每个组播组都设置一个计时器，计时器的值在 0～最大响应时间中随机选取，当其中任何一个计时器的值减为 0 时，主机就会发送该组播组的成员报告消息。

合理设置最大响应时间，可以使主机快速响应查询信息，路由器也就能快速地掌握组播组成员的存在状况。

请在接口视图下进行下列配置。

操作	命令
配置 IGMP 最大查询响应时间	**igmp max-response-time** *seconds*

操作	命令
恢复缺省值	**undo igmp max-response-time**

最大响应时间的值愈小，路由器阻断组的速度愈快。实际中响应时间是设置值范围（1～25 秒）内的一个随机值。

缺省情况下，IGMP 最大查询响应时间为 10 秒。

3.2.6 配置路由器加入指定组播组

通常情况下，运行 IGMP 的主机会对组播路由器的 IGMP 查询报文进行响应，如果由于某种原因无法响应，就可能导致组播路由器认为该网段没有该组播组的成员，从而取消相应的路径。

为避免这种情况的发生，可以将路由器的某个接口配置成为组播组成员，当从该接口收到 IGMP 查询报文时，由路由器进行响应，从而保证接口所在网段能够继续收到组播报文。

请在接口视图下进行下列配置。

操作	命令
配置接口加入组播组	**igmp host-join** *group-address*
取消接口作为组成员	**undo igmp host-join** *group-address*

缺省情况下，路由器不加入任何组播组。

3.2.7 控制接口对组播组的访问

组播路由器通过接收到的 IGMP 组成员关系报文来确定该网段的组成员关系。可以通过此项配置在每一个接口上设一个过滤器，以限制接口所服务的组播组范围。

请在接口视图下进行下列配置。

操作	命令
控制接口对组播组的访问	**igmp group-policy** *acl-number* [**1** | **2**]
取消接口对组播组的访问控制	**undo igmp group-policy**

这里参数 **1** 和 **2** 是指 IGMP 的版本号。

缺省不设置过滤器，即接口允许任一组播组。

3.2.8 启用 IGMP 组快速离开功能

本任务用于启用接口的 IGMP 组快速离开功能。执行该配置后，当接口收到用户的 IGMP leave 报文后，不再发送 IGMP 组成员查询报文，而是直接将用户从 IGMP 组中删除。

请在接口视图下进行下列配置。

操作	命令
启用接口的 IGMP 组快速离开功能	**igmp fast-leave**
关闭接口的 IGMP 组快速离开功能	**undo igmp fast-leave**

缺省情况下，接口未启用 IGMP 组快速离开功能。

3.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后 IGMP 的运行情况，验证配置的效果。

在用户视图下，执行 **debugging** 命令可对 IGMP 进行调试。

操作	命令
显示 IGMP 组播组成员信息	**display igmp group** [**local** | *group-address* | **interface** *interface-type interface-number*]
显示接口的 IGMP 配置和运行信息	**display igmp interface** [*interface-type interface-number*]
显示用户的 IGMP 配置和运行信息	**display igmp user** { **user-id** *id* | **ip-address** *ip-address* }
打开 IGMP 调试信息开关	**debugging igmp** { **all** | **event** | **host** | **packet** | **timer** }
关闭 IGMP 调试信息开关	**undo debugging igmp** { **all** | **event** | **host** | **packet** | **timer** }
打开接口板的 IGMP 调试信息开关	**debugging lpu** *slot* **igmp** { **all** | **event** | **host** | **packet** | **timer** | **user** }
关闭接口板的 IGMP 调试信息开关	**undo debugging lpu** *slot* **igmp** { **all** | **event** | **host** | **packet** | **timer** | **user** }

3.4 故障排除

3.4.1 组播转发异常

1. 故障现象

IGMP 主机从接口加入，PIM-SM（或 PIM-DM）却未能将该接口加入转发项。

2. 故障排除

(1) 检查该接口的状态是否为 UP。
(2) 检查接口上是否配置 IP 地址。
(3) 检查接口是否启动使能 PIM-SM（或 PIM-DM）。

3.4.2 IGMP 组成员关系不一致

1. 故障现象

在同一网段的不同路由器之间，IGMP 组成员关系不一致。

2. 故障排除

分别在同一网段所有路由器上执行命令 **display igmp interface**，检查 IGMP 所有参数，确保连接在同一网段的所有路由器接口的 IGMP 所有参数完全一致。

第4章 配置 PIM-DM

本章包括以下内容：

- PIM-DM 简介
- 配置内容
- 显示和调试
- 配置实例

4.1 PIM-DM 简介

有关 PIM-DM 的基本概念请参考“1.3.2 1. PIM-DM”。

4.1.1 PIM-DM 工作过程

PIM-DM 的工作过程可以概括为邻居发现、扩散-剪枝（Flooding & Prune）、嫁接（Graft）三个阶段。

1. 邻居发现

PIM-DM 路由器刚开始启动时，需要使用 Hello 报文来进行邻居发现。各个运行 PIM-DM 的网络节点之间使用 Hello 报文保持相互之间的联系。PIM-DM 的 Hello 报文是周期性发送的。

2. 扩散-剪枝过程

扩散-剪枝过程简单描述如下：

- PIM-DM 首先假设网络上的所有主机都准备接收组播数据。
- 当某组播源 S 开始向组播组 G 发送数据时，路由器接收到组播报文后，首先根据单播路由表进行 RPF 检查。
- 如果检查通过，路由器创建一个（S，G）表项，然后将数据向网络上所有下游 PIM-DM 节点转发（Flooding）。
- 如果没有通过 RPF 检查，即组播报文从错误的接口输入，则将报文丢弃。经过这个过程，在 PIM-DM 组播域内，都会创建一个（S，G）表项。
- 如果下游节点没有组播组成员，则向上游节点发剪枝（Prune）消息，通知上游节点不用再向下游节点转发数据。
- 上游节点收到剪枝消息后，就将相应的接口从其组播转发表项（S，G）对应的输出接口列表中删除。

以上过程就称为扩散-剪枝过程。各个被剪枝的节点同时提供了超时机制，当剪枝超时的时候，每台路由器又重新开始扩散-剪枝过程。PIM-DM 的扩散-剪枝机制周期性地不断进行。

3. 嫁接

当被剪枝的下游节点需要恢复到转发状态时，该节点使用嫁接报文通知上游节点。

4.1.2 Assert 机制

如图所示，如果处于一个 LAN 网段上的两台组播路由器 A 和 B，都各自有到组播源 S 的接收途径，那么它们在接收到组播源 S 发出的组播数据报文以后，都会向 LAN 上转发该组播报文，这时候下游节点组播路由器 C 就会收到两份相同的组播报文。

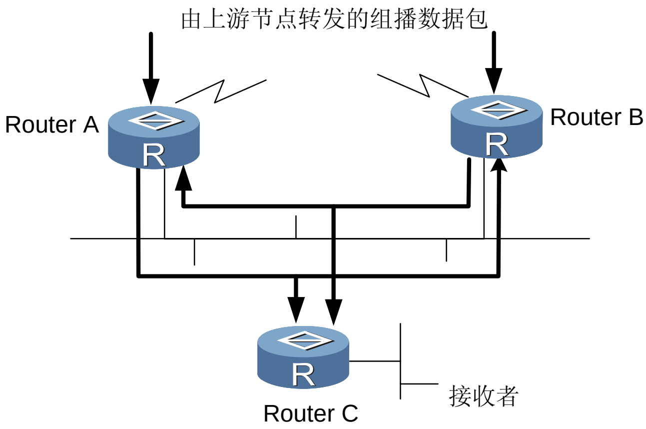


图4-1 Assert 机制示意图

路由器检测到这种情况发生以后，就需要通过 Assert 机制来选定一个唯一的转发者。通过发送 Assert 报文，选出一个最优的路径，如果两条或两条以上路径的优先级和 metric 相同，则选择 IP 地址最大的作为该（S，G）项的上游邻居，由它负责该（S，G）组播报文的转发。

4.1.3 RPF 检查

有关 RPF 检查，请参见“1.4.1 逆向路径转发 RPF”。

4.2 配置内容

配置 PIM-DM 需要依次进行如下操作，当路由器在 PIM-DM 协议域中运行时，建议在非边界路由器的全部接口启动 PIM-DM。

PIM-DM 配置包括以下内容：

- 启动组播

- 启动 PIM-DM 协议
- 配置接口的 Hello 报文发送间隔

4.2.1 启动组播

请参见“2.1.1 启动组播”。

4.2.2 启动 PIM-DM 协议

PIM-DM 协议需要分别在各个接口配置中启动。

在接口上配置了 PIM-DM 之后，协议会定期发送 PIM-DM 协议 Hello 报文，并且处理 PIM-DM 邻居发送的协议报文。

请在接口视图下进行下列配置。

操作	命令
在接口上启动 PIM-DM 协议	**pim dm**
在接口上禁止 PIM-DM 协议	**undo pim dm**

通常情况下，建议各个接口全部配置 PIM-DM。此配置必须在系统视图下使能组播路由之后，才能生效。

在接口上启动了 PIM-DM 协议后，不能再对此接口启动 PIM-SM 协议，反之亦然。

4.2.3 配置接口的 Hello 报文发送间隔

接口启动 PIM 协议后，会定期发送 Hello 报文。发送 Hello 报文的间隔可以依据与接口相连网络的带宽和类型适当加以修改。

请在接口视图下进行下列配置。

操作	命令
设置接口的 Hello 报文发送间隔	**pim timer hello** *seconds*
恢复该时间间隔的缺省值	**undo pim timer hello**

缺省的 Hello 报文发送间隔是 30 秒。用户可以根据不同的网络环境来配置。一般情况下，无需修改该参数。

该配置必须在接口视图下启动 PIM 协议（PIM-DM 或者 PIM-SM）后才允许配置。

4.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后 PIM-DM 的运行情况，验证配置的效果。

在用户视图下，执行 **debugging** 命令可对 PIM-DM 进行调试。

操作	命令
显示 PIM 协议组播路由表	**display pim routing-table** [{ { ***g** [*group-address* [**mask** { *mask-length* | *mask* }]] | ****rp** [*rp-address* [**mask** { *mask-length* | *mask* }]] } | { *group-address* [**mask** { *mask-length* | *mask* }] | *source-address* [**mask** { *mask-length* | *mask* }] } * } | **incoming-interface** { *interface-type interface-num* | *interface-name* | **null** } | { **dense-mode | sparse-mode** }] *
显示 PIM 协议接口信息	**display pim interface** [*interface-type interface-number*]
显示 PIM 相邻路由器信息	**display pim neighbor** [**interface** *interface-type interface-number*]
打开 PIM 协议调试开关	**debugging pim common** { **all** | **event** | **packet** | **timer** }
关闭 PIM 协议调试开关	**undo debugging pim common** { **all** | **event** | **packet** | **timer** }
打开 PIM-DM 协议调试开关	**debugging pim dm** { **all** | **mrt** | **alert** | { **recv** | **send** } { **all** | **assert** | **graft** | **graft-ack** | **join** | **prune** } }
关闭 PIM-DM 协议调试开关	**undo debugging pim dm** { **all** | **mrt** | **alert** | { **recv** | **send** } { **all** | **assert** | **graft** | **graft-ack** | **join** | **prune** } }

4.4 配置实例

1. 组网需求

在图 4-2中，Multicast Source 作为组播源，Receiver 1 和 Receiver 2 是该组播组的两个接收成员。

2. 组网图

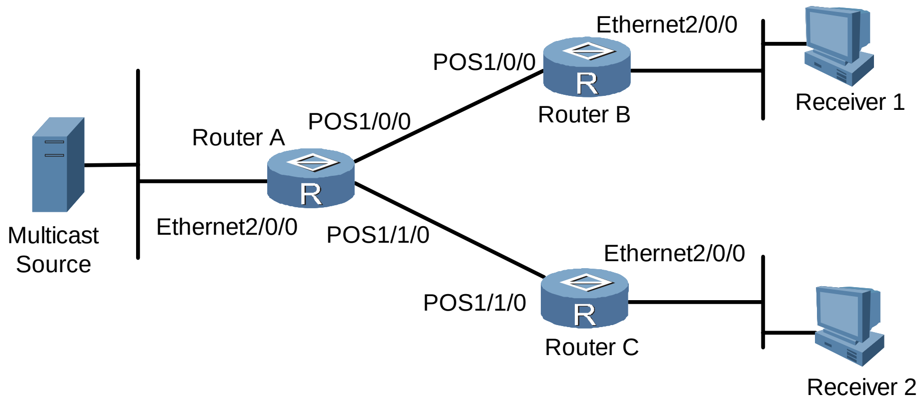


图4-2 PIM-DM 典型配置组网图

3. 配置步骤

(1) 配置 Router A

启动组播。

```
<Quidway> system-view
[Quidway] multicast routing-enable
```

分别在接口 Ethernet2/0/0、POS1/0/0 和 POS1/1/0 上启动 PIM-DM。

```
[Quidway] interface pos 1/0/0
[Quidway-Pos1/0/0] pim dm
[Quidway-Pos1/0/0] quit
[Quidway] interface pos 1/1/0
[Quidway-Pos1/1/0] pim dm
[Quidway-Pos1/1/0] quit
[Quidway] interface ethernet 2/0/0
[Quidway-Ethernet2/0/0] pim dm
```

(2) 配置 Router B

启动组播。

```
<Quidway> system-view
[Quidway] multicast routing-enable
```

在接口 POS1/0/0 上启动 PIM-DM。

```
[Quidway] interface pos 1/0/0
[Quidway-Pos1/0/0] pim dm
[Quidway-Pos1/0/0] quit
```

在接口 Ethernet2/0/0 上启动 PIM-DM。

```
[Quidway] interface ethernet 2/0/0
```

```
[Quidway-Ethernet2/0/0] igmp enable
[Quidway-Ethernet2/0/0] pim dm
```

(3) 配置 Router C

启动组播。

```
<Quidway> system-view
[Quidway] multicast routing-enable
```

在接口 POS1/1/0 和 Ethernet2/0/0 上启动 PIM-DM。

```
[Quidway] interface pos 1/1/0
[Quidway-Pos1/1/0] pim dm
[Quidway-Pos1/1/0] quit
[Quidway] interface ethernet 2/0/0
[Quidway-Ethernet2/0/0] igmp enable
[Quidway-Ethernet2/0/0] pim dm
```

第5章 配置 PIM-SM

本章包括以下内容：

- PIM-SM 简介
- 配置内容
- 显示和调试
- 配置实例
- 故障排除

5.1 PIM-SM 简介

有关 PIM-SM 的基本概念请参考“1.3.2 2. PIM-SM”。

5.1.1 PIM-SM 工作过程

PIM-SM 的工作过程可以概括为邻居发现、RPT（RP Shared Tree）的生成、组播源注册、SPT 切换四个阶段。

1. 邻居发现

请参见“4.1.1 1. 邻居发现”。

2. RPT 的生成

RPT 的生成过程如图 5-1所示。

- 当主机加入一个组播组 G 时，与该主机直接相连的叶子路由器通过 IGMP 报文了解到有组播组 G 的接收者。
- 叶子路由器为组播组 G 计算出对应的汇聚点 RP，然后向 RP 方向的上一级节点发送加入组播组的消息（join 消息）。
- 从叶子路由器到 RP 之间途经的每个路由器都会在转发表中生成（*，G）表项，表示无论是哪个源发出的，发送至组播组 G 的，都适用于该表项。
- 当 RP 收到发往组播组 G 的报文后，报文就会沿着已经建立好的路径到达叶子路由器，进而到达主机。这样就生成了以 RP 为根的 RPT。

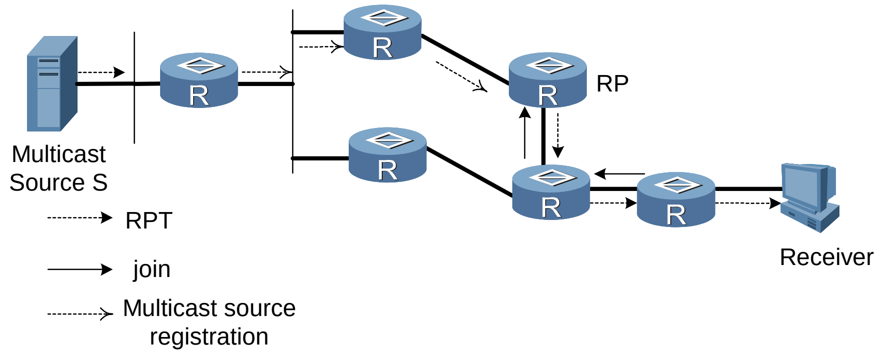


图5-1 RPT 示意图

3. 组播源注册

当组播源 S 向组播组 G 发送了一个组播报文时，与 S 直接相连的 PIM-SM 组播路由器负责将接收到的该组播报文封装成注册报文，以单播形式发给对应的 RP。如果一个网段上有多个 PIM-SM 组播路由器，将由指定路由器 DR 发送该组播报文。

4. SPT 切换

当组播路由器发现从 RP 发来的目的地址为 G 的组播报文的速率超过了阈值时，组播路由器向源 S 的上一级节点发送加入消息，导致 RPT 向 SPT 的切换。

5.1.2 PIM-SM 配置前准备工作

1. 配置候选 RP

在 PIM-SM 网络中，可以存在多个候选 RP（Candidate-RP），每个候选 RP 负责转发目的地址在一定范围内的组播报文。配置多个候选 RP 可以实现 RP 负载分担。候选 RP 之间没有主次之分，所有的组播路由器收到 BSR（Boot Strap Router）通告的候选 RP 消息后，根据相同的算法计算出与某一组播组对应的 RP。

说明：

一个 RP 可以为多个组播组服务，也可以为所有组播组服务。每个组播组在任意时刻，只能唯一地对应一个 RP，不能同时对应多个 RP。

2. 配置 BSR

BSR 是 PIM-SM 网络里的管理核心，它负责收集候选 RP 发来的信息，并把它们广播出去。

一个网络内部只能有一个 BSR，但可以配置多个候选 BSR（Candidate-BSR）。这样，一旦某个 BSR 发生故障后，能够切换到另外一个。候选 BSR 通过自动选举产生 BSR。

3. 配置静态 RP

RP 是组播路由中的核心路由器，如果由于某种原因使由 BSR 机制选举产生的动态 RP 失效，则可以通过配置静态 RP 来指定 RP。静态 RP 作为动态 RP 的备份，可以提高网络的健壮性，增强组播网络的运营管理能力。

5.2 配置内容

配置 PIM-SM 需要进行如下操作，当路由器在 PIM-SM 协议域中运行时，建议在非边界路由器的全部接口启动 PIM-SM。

PIM-SM 配置包括以下内容：

- 启动组播
- 启动 PIM-SM 协议
- 配置接口的 Hello 报文发送间隔
- 设置 PIM-SM 域边界
- 进入 PIM 视图
- 配置候选 BSR
- 配置候选 RP
- 配置 RP 过滤 DR 发送的注册报文
- 配置从 RPT 切换到 SPT 的阈值

其中，前四项配置是必选的，其他五项可使用缺省配置。

说明：

在整个 PIM-SM 域中，至少要在一台路由器上配置候选 RP 和候选 BSR。

5.2.1 启动组播

请参见“2.1.1 启动组播”。

5.2.2 启动 PIM-SM 协议

PIM-SM 协议需要分别在各个接口配置中启动。

在接口上配置了 PIM-SM 之后，协议会定期发送 PIM 协议 Hello 报文，并且处理 PIM 邻居发送的协议报文。

请在接口视图下进行下列配置。

操作	命令
在接口上启动 PIM-SM 协议	**pim sm**
在接口上禁止 PIM-SM 协议	**undo pim sm**

通常情况下，建议各个接口全部配置 PIM-SM。此配置必须在系统视图下使能组播路由之后，才能生效。

在接口上启动了 PIM-SM 协议后，不能再对此接口启动 PIM-DM 协议，反之亦然。

5.2.3 配置接口的 Hello 报文发送间隔

接口启动 PIM 协议后，会定期发送 Hello 报文。发送 Hello 报文的间隔可以依据与接口相连网络的带宽和类型适当加以修改。

请在接口视图下进行下列配置。

操作	命令
设置接口的 Hello 报文发送间隔	**pim timer hello** *seconds*
恢复该时间间隔的缺省值	**undo pim timer hello**

缺省的 Hello 报文发送间隔是 30 秒。用户可以根据不同的网络环境来配置。一般情况下，无需修改该参数。

该配置必须在接口视图下启动 PIM 协议（PIM-DM 或者 PIM-SM）后才允许配置。

5.2.4 设置 PIM-SM 域边界

设置 PIM-SM 域边界后，自举报文（Bootstrap message）从任何方向都不能穿过边界。通过这种方法，可以分割 PIM-SM 域。

请在接口视图下进行下列配置。

操作	命令
设置 PIM-SM 域边界	**pim bsr-boundary**
删除配置的 PIM-SM 域边界	**undo pim bsr-boundary**

缺省时不设置域边界。在缺省配置下，自举报文不能通过该边界，但其他 PIM 报文可以通过域边界。缺省配置能够有效地将网络划分成使用不同 BSR 的区域。

5.2.5 进入 PIM 视图

配置与 PIM 相关的全局参数，如候选 BSR、候选 RP 等需要进入 PIM 视图。

请在系统视图下进行下列配置。

操作	命令
进入 PIM 视图	**pim**
退回系统视图	**undo pim**

使用 **undo pim** 命令将清除在 PIM 视图下进行的配置。

5.2.6 配置候选 BSR

在一个 PIM 域中有一个 BSR 负责收集并发布 RP 信息。为了产生自举路由器 BSR，需要配置一个或多个候选 BSR，候选 BSR 之间通过自动选举。下面简单描述一下候选 BSR 之间的自动选举：

将路由器配置为候选 BSR 的同时，必须指定一个启动了 PIM-SM 的接口。

最初，每个候选 BSR 都认为自己是本 PIM-SM 的 BSR，并使用这个接口的 IP 地址作为 BSR 地址，发送自举报文。

当候选 BSR 收到其他路由器发来的自举报文时，它将新收到的自举报文的 BSR 地址与自己的 BSR 地址进行比较，比较标准包括优先级和 IP 地址，优先级相同的情况下，较大的 IP 地址被认为是更好的。如果前者更好，则将这个新的 BSR 地址替换自己的 BSR 地址，并且不再认为自己是 BSR；否则，保留自己的 BSR 地址，继续将自己视为 BSR。

请在 PIM 视图下进行下列配置。

操作	命令
配置候选 BSR	**c-bsr** *interface-type interface-number mask-len* [*priority*]
取消候选 BSR 的配置	**undo c-bsr**

候选 BSR 应配在骨干网的路由器上。缺省没有设置 BSR，优先级缺省值为 0。

注意：

在同一台路由器上只能配置一个候选 BSR，在另外的接口上配置候选 BSR 会取代先前的配置。

5.2.7 配置候选 RP

在 PIM-SM 协议中，路由组播数据创建的共享树是以 RP 为树根的，从组播组到 RP 存在一个映射，一个组播组映射到一个 RP 上，不同的组可以映射到同一 RP 上。

请在 PIM 视图下进行下列配置。

操作	命令
配置候选 RP	**c-rp** *interface-type interface-number* [**group-policy** *acl-number*] [**priority** *priority-value*]
取消候选 RP 的配置	**undo c-rp** *interface-type interface-number*

配置 RP 时，若没有指定所服务的组播组范围，该 RP 为所有组播组服务，否则，所服务组的范围限定为指定范围的组播组。

建议在骨干网路由器上配置候选 RP。

5.2.8 配置 RP 过滤 DR 发送的注册报文

在 PIM-SM 网络中，通过注册报文过滤机制，可以在 RP 上控制哪些源可以向哪些组发送报文，即 RP 可以过滤 DR 发送来的注册报文，只接受特定的报文。

请在 PIM 视图下进行下列配置。

操作	命令
配置 RP 过滤 DR 发送的注册报文	**register-policy** *acl-number*
取消 RP 过滤 DR 发送的注册报文	**undo register-policy**

如果某个源组表项（S，G）被 ACL 拒绝，或者 ACL 没有定义对它的操作，或者没有定义 ACL，RP 都会向 DR 发送 RegisterStop 消息，阻止该组播数据流的注册过程。

⚠ **注意：**

只有与 ACL 的 permit 语句匹配的注册报文才会被 RP 接受。指定一个没有定义的 ACL 将使 RP 拒绝所有的注册报文。

5.2.9 配置从 RPT 切换到 SPT 的阈值

PIM-SM 路由器最初通过共享树（RPT）转发组播数据包，但是如果组播数据通过的速率超过一定的阈值，数据包经过的最后一跳路由器则会发起从共享树到最短路径树（SPT）的切换。

请在 PIM 视图下进行下列配置。

操作	命令
配置从 RPT 切换到 SPT 的阈值	**spt-switch-threshold** { *traffic-rate* | **infinity** } [**group-policy** *acl-number*]
恢复缺省值	**undo spt-switch-threshold** { *traffic-rate* | **infinity** } [**group-policy** *acl-number*]

缺省情况下，阈值为 0，即最后一跳路由器收到第一个组播数据包就会转向最短路径树。

5.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示 PIM-SM 的运行情况，验证配置的效果。

在用户视图下，执行 **debugging** 命令可对 PIM-SM 进行调试。

操作	命令
显示 BSR 信息	**display pim bsr-info**
显示 RP 信息	**display pim rp-info** [*group-address*]
打开 PIM-SM 协议调试开关	**debugging pim sm** { **all** | **mbr** { **alert** | **fresh** } | **mrt** | **msdp** | **recv** { **assert** | **bootstrap** | **crpadv** | **jp** | **reg** | **regstop** } | **send** { **assert** | **bootstrap** | **crpadv** | **jp** | **reg** | **regstop** } | **timer** { **assert** | **bsr** | **crpadv** | **jp** | **jpdelay** | **mrt** | **probe** | **spt** } | **verbose** | **warning** }
关闭 PIM-SM 协议调试开关	**undo debugging pim sm** { **all** | **mbr** { **alert** | **fresh** } | **mrt** | **msdp** | **recv** { **assert** | **bootstrap** | **crpadv** | **jp** | **reg** | **regstop** } | **send** { **assert** | **bootstrap** | **crpadv** | **jp** | **reg** | **regstop** } | **timer** { **assert** | **bsr** | **crpadv** | **jp** | **jpdelay** | **mrt** | **probe** | **spt** } | **verbose** | **warning** }

5.4 配置实例

1. 组网需求

在实际的网络中，由于路由设备由不同的厂商提供，设备上的路由协议也会各不相同。但由于 PIM 协议是独立于特定的单播路由协议的，所以本节不关心单播路由协议，假定各个路由器之间相互可达。

2. 组网图

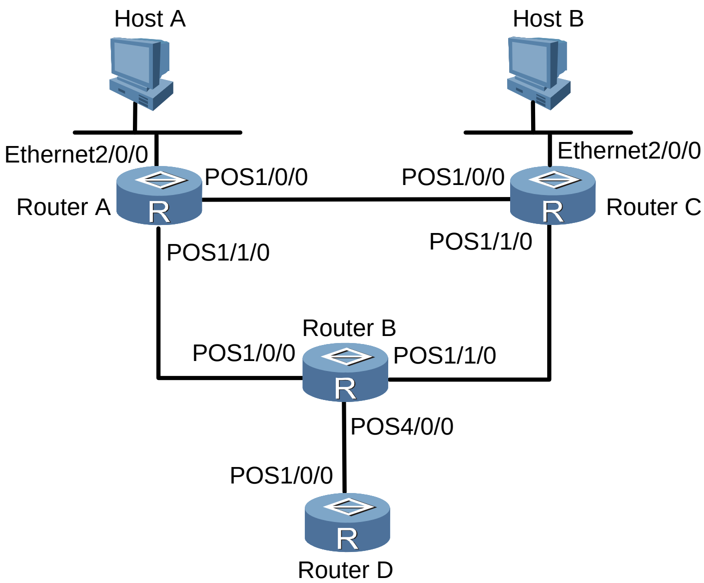


图5-2 PIM-SM 综合配置图

3. 配置步骤

(1) 配置 Router A

启动 PIM-SM。

```
<Router A> system-view
[Router A] multicast routing-enable
[Router A] interface ethernet 2/0/0
[Router A-Ethernet2/0/0] igmp enable
[Router A-Ethernet2/0/0] pim sm
[Router A-Ethernet2/0/0] quit
[Router A] interface pos 1/0/0
[Router A-Pos1/0/0] pim sm
[Router A-Pos1/0/0] quit
[Router A] interface pos 1/1/0
[Router A-Pos1/1/0] pim sm
```

配置特定组播组从共享树切换到最短路径树的阈值为 10kbit/s。

```
[Router A] acl number 2004
[Router A-acl-basic-2004] rule permit source 225.0.0.0 0.0.0.255
[Router A-acl-basic-2004] quit
[Router A] pim
[Router A-pim] spt-switch-threshold 10 group-policy 2004
```

(2) 配置 Router C

启动 PIM-SM。

```
<Router C> system-view
[Router C] multicast routing-enable
[Router C] interface ethernet 2/0/0
[Router C-Ethernet2/0/0] pim sm
[Router C-Ethernet2/0/0] quit
[Router C] interface pos 1/0/0
[Router C-Pos1/0/0] pim sm
[Router C-Pos1/0/0] quit
[Router C] interface pos 1/1/0
[Router C-Pos1/1/0] pim sm
```

(3) 配置 Router B

启动 PIM-SM。

```
<Router B> system-view
[Router B] multicast routing-enable
[Router B] interface pos 1/0/0
[Router B-Pos1/0/0] pim sm
[Router B-Pos1/0/0] quit
[Router B] interface pos 1/1/0
[Router B-Pos1/1/0] pim sm
[Router B-Pos1/1/0] quit
[Router B] interface pos 4/0/0
[Router B-Pos4/0/0] pim sm
[Router B-Pos4/0/0] quit
```

配置候选 BSR。

```
[Router B] pim
[Router B-pim] c-bsr pos 1/0/0 30 2
[Router B-pim] quit
```

配置候选 RP。

```
[Router B] acl number 2004
[Router B-acl-basic-2004] rule permit source 225.0.0.0 0.255.255.255
[Router B-acl-basic-2004] quit
```

```
[Router B] pim
[Router B-pim] c-rp pos 1/0/0 group-policy 2004
[Router B-pim] quit
```

配置 PIM 边界。

```
[Router B] interface pos4/0/0
[Router B-Pos4/0/0] pim bsr-boundary
```

在接口 POS4/0/0 上配置了边界后，路由器 Router D 就收不到路由器 Router B 发出的 BSR 信息，它被排除在本 PIM 域之外。

假设主机 A 是 225.0.0.1 的接收者，主机 B 现在开始发送目的地址为 225.0.0.1 的数据。开始时，路由器 A 通过路由器 B 接收主机 B 发送的组播数据，当主机 B 发送的组播数据的速率超过 10kbit/s 后，路由器 A 会加入最短路径树，直接从路由器 C 接收主机 B 发送的组播数据包。

5.5 故障排除

1. 故障现象

路由器不能正确建立组播路由表。

2. 故障排除

可以按照如下步骤进行：

(1) 解决问题之前，首先应保证单播路由是正确的。

(2) PIM-SM 协议需要有 RP 和 BSR 的支持，所以首先使用 **display pim bsr-info**，看是否有 BSR 信息，如果不存在，则需要查看是否有通向 BSR 的单播路由。这个问题解决后，看 RP 信息是否正确，使用 **display pim rp-info** 命令。如果没有 RP 信息，也需要检查单播路由。

(3) 使用 **display pim neighbor** 来查看是否正确建立了邻居关系。

第6章 配置 MSDP

本章包括以下内容：

- MSDP 简介
- 配置内容
- 显示和调试
- 配置实例

6.1 MSDP 简介

6.1.1 MSDP 概述

MSDP 用来发现其他 PIM-SM 域内的组播源信息。配置了 MSDP 对等体的 RP 将其域内的活动组播源信息通过 SA（Source Active）消息通告给它的所有 MSDP 对等体。由此一个 PIM-SM 域内的组播源信息就会被传递到另一个 PIM-SM 域。

MSDP 对等体可以建立在不同域的 RP 之间，也可以建立在同一域的多个 RP 之间，还可以建立在 RP 与普通路由器之间或者普通路由之间。MSDP 对等体之间使用 TCP 连接。

MSDP 使得一个 PIM-SM 域不需要依赖另一个 PIM-SM 域内的 RP，因为在得到另一个 PIM-SM 域内的组播源信息之后，一个 PIM-SM 域里的接收者可以不通过另一 PIM-SM 域里的 RP 而直接加入到这个域内组播源的 SPT 上。

另外，MSDP 通过 RPF 检查机制，只接受从正确路径上接收到的 SA 消息，避免接受冗余的 SA 消息；可以通过配置 Mesh 全连接组来避免 SA 消息在 MSDP 对等体之间泛滥。

6.1.2 Anycast RP

MSDP 另一个应用是 Anycast RP。在一个域内，用同一个 IP 地址配置不同的路由器上的某一接口（通常是 Loopback 接口），同时，配置这些路由器上这个接口为候选 RP，并在这些 RP 之间建立 MSDP 对等体关系。

单播路由收敛后，组播源可以选择最近的 RP 注册，接收者也可以选择最近的 RP 加入其 RPT。这些 RP 之间通过 MSDP 对等体了解对方的注册源信息，最终每个 RP 了解到整个域内的所有组播源，这样，每个 RP 上的接收者就可以接收到整个域内的所有组播源发出的组播数据。

组播源通过向就近的 RP 发起注册和 RPT 加入，可以实现 RP 的负载分担；一个 RP 失效后，其原来注册的源和加入者，又会选择另一个就近的 RP 注册和加入，实现了 RP 的冗余备份。

6.1.3 MSDP 工作原理

1. 识别组播源和接收组播数据过程

如图 6-1所示，PIM-SM 域 1 和域 2、域 3 中的 RP 之间建立 MSDP 对等体关系。域 3 中存在某个组的成员。

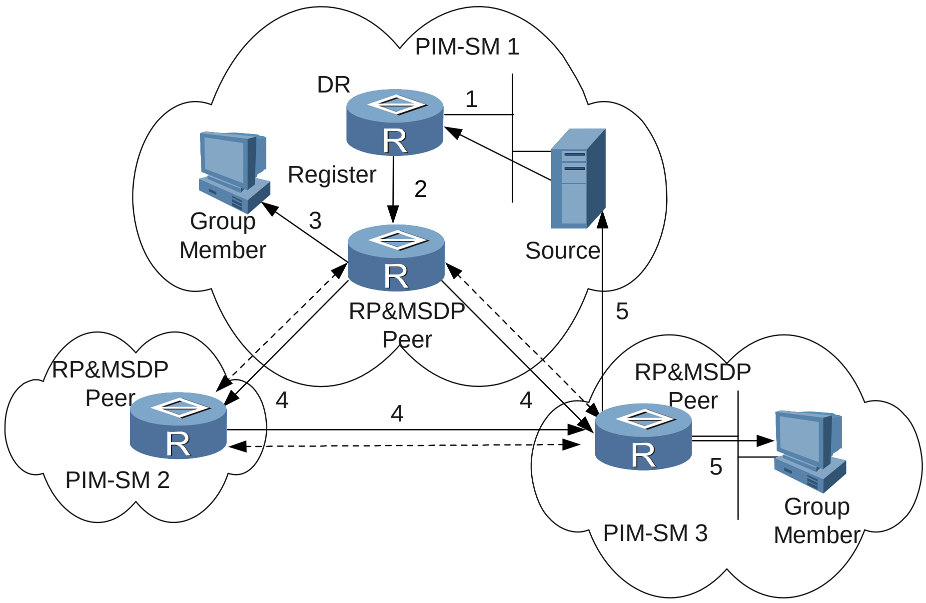


图6-1 MSDP 的工作原理图（一）

当域 1 中的一个组播源向这个组发送数据时，域 3 中的组成员从得知这个组播源到接收到这个组播源的组播数据的工作过程如下：

(1) PIM-SM 域 1 里的组播源开始发送数据包。

(2) 连接组播源的指定路由器 DR 将组播源发出的数据封装在 Register 报文里，发给域 1 内的 RP。

(3) 域 1 中的 RP 将报文解封装，沿域内的共享树向下转发给域内的所有成员，域内成员可以选择是否切换到源树上。

(4) 同时，域 1 内的 RP 将生成一个 SA 消息，发送给 MSDP 对等体（PIM-SM 域 2 和域 3 内的 RP）。

说明：

SA 消息包括如下内容：组播源的 IP 地址、组播组地址和生成消息的 RP 的地址，并且域 1 内的 RP 也将收到的第一个组播数据封装在这个 SA 消息中。

(5) 如果在 MSDP 对等体所在的域（PIM-SM 域 3）里有组成员，则该域内的 RP 将 SA 消息中封装的组播数据沿共享树下发到组成员的同时，向组播源发送加入消息。

(6) 当逆向转发路径建立起来之后，组播源发出的数据将直接发送到域 3 的 RP 上，RP 向共享树转发数据。此时，域 3 中连接组成员的最后一跳路由器可以选择是否切换到 SPT 上。

2. MSDP 对等体之间消息转发过程和 RPF 检查过程

如图 6-2所示，Router A、Router B、Router C、Router D、Router E、Router F 分别属于域 1、域 2 和域 3，并在相互之间建立 MSDP 对等体关系，在图中用双向的箭头表示。其中 Router B、Router C、Router D 之间建立 Mesh 全连接组。在 MSDP 对等体之间转发 SA 消息的过程和 RPF 检查的方法如下：

(1) 如果发出 SA 消息的 MSDP 对等体就是组播源所在的 RP，接受该 SA 消息，并向其他对等体转发，如：Router A 发给 Router B；

(2) 如果只配置了一个 MSDP 对等体，接受该对等体发来的 SA 消息，如 Router B 发给 Router A；

(3) 如果 SA 消息是从静态 RPF 对等体发来的，接受该 SA 消息，并向其他对等体转发，如 Router D 发给 Router E；

(4) 如果 SA 消息是从 MSDP 全连接组中的对等体发来的，接受该 SA 消息，并向该全连接组以外的对等体转发，如 Router B 发给 Router D；

(5) 如果 SA 消息是从同一个域内的 MSDP 对等体发来的，并且该对等体是到源所在域 RP 的最佳路径上的下一跳，接受该 SA 消息，并向其他对等体转发，如 Router E 发给 Router F；

(6) 如果 SA 消息是从不同域的 MSDP 对等体发来的，并且它是到源所在域 RP 最佳路径上的下一个自治域，接受该 SA 消息，并向其他对等体转发，如：Router D 发给 Router F；

(7) 其他 SA 消息，不接受也不转发。

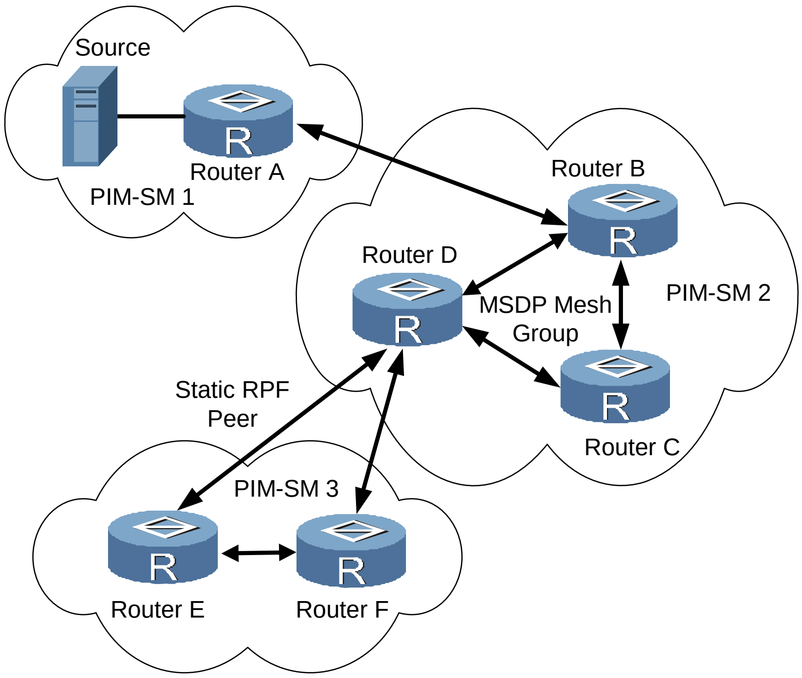


图6-2 MSDP 的工作原理图（二）

说明：

配置 MSDP 的路由器同时需要运行 BGP 或 MBGP，并且，建议 MSDP 对等体的地址与 BGP 或 MBGP 对等体的地址相同。如果不运行 BGP 或 MBGP，需要配置静态 RPF 对等体。

6.2 配置内容

配置 MSDP 包括以下内容：

- 使能 MSDP
- 配置 MSDP 对等体
- 配置静态 RPF 对等体
- 配置 Originating RP
- 配置缓存 SA 状态
- 配置缓存 SA 的最大数量
- 配置请求源信息
- 配置对源信息的处理
- 配置 MSDP 全连接组
- 配置 MSDP 连接重试周期

- 关闭 MSDP 对等体
- 复位 MSDP

使能 MSDP 和配置 MSDP 对等体是必须的，其余则是可选的。

6.2.1 使能 MSDP

在配置 MSDP 前，必须先使能 MSDP。

请在系统视图下进行下列配置。

操作	命令
使能 MSDP 并进入 MSDP 视图	**msdp**
清除 MSDP 的所有配置	**undo msdp**

6.2.2 配置 MSDP 对等体

使能 MSDP 后，还需要在本地配置 MSDP 对等体。

请在 MSDP 视图进行下列配置。

操作	命令
配置 MSDP 对等体	**peer** *peer-address* **connect-interface** *interface-type interface-number*
取消配置 MSDP 对等体	**undo peer** *peer-address*
为 MSDP 对等体添加描述性文本	**peer** *peer-address* **description** *text*
删除已添加的描述性文本	**undo peer** *peer-address* **description**

如果本地路由器与一个 MSDP 对等体同时也是 BGP 对等体，MSDP 对等体和 BGP 对等体应使用相同的 IP 地址。

存在 MSDP 对等体关系的两台路由器上不一定需要运行 BGP 或 MBGP，只要它们之间存在一条 BGP 或 MBGP 路径就可以。而如果没有 BGP 或 MBGP 路径，则必须配置静态 RPF 对等体。

6.2.3 配置静态 RPF 对等体

请在 MSDP 视图进行下列配置。

操作	命令
配置静态 RPF 对等体	**static-rpf-peer** *peer-address* [**rp-policy** *list*]
取消配置的静态 RPF 对等体	**undo static-rpf-peer** *peer-address*

缺省情况下，没有静态 RPF 对等体。

必须先配置 **peer** 命令，然后才能配置 **static-rpf-peer**。并且，如果命令 **peer** 只配置了一个 MSDP 对等体，这个 MSDP 对等体将被作为静态 RPF 对等体。

当同时配置多于一个的静态 RPF 对等体时，必须遵守如下两种配置方法：

- 都使用 **rp-policy** 参数

多个静态 RPF 对等体同时起作用，并按照所配置的前缀列表对 SA 消息中的 RP 进行过滤，只接收 RP 地址通过过滤的 SA 消息。如果同时配置了多个使用相同 **rp-policy** 参数的静态 RPF 对等体，则从其中一个对等体接收到的 SA 消息会向其他对等体转发。

- 都不使用 **rp-policy** 参数

按照配置的先后顺序，只有第一个连接状态是 UP 的静态 RPF 对等体是激活的，接收来自该对等体所有的 SA 消息，从其他静态 RPF 对等体接收到的 SA 消息将被丢弃。如果这个激活的静态 RPF 对等体失效（如配置取消或连接断开），仍然按照配置的先后顺序，重新选择第一个连接状态是 UP 的静态 RPF 对等体作为激活的静态 RPF 对等体。

6.2.4 配置 Originating RP

创建 SA 消息时，可以配置 MSDP 使用指定接口的 IP 地址作为其 SA 消息中的 RP 地址。

请在 MSDP 视图进行下列配置。

操作	命令
使用指定接口的 IP 地址作 RP 地址	**originating-rp** *interface-type interface-number*
取消上述操作	**undo originating-rp**

缺省情况下，SA 消息的 RP 地址为 PIM 配置的 RP 地址。

6.2.5 配置缓存 SA 状态

在路由器上配置缓存 SA 状态，可以使后续新加入的组不必等待下一次到达的 SA 消息，就可以从 SA 缓存中获得所有的活动源，加入到相应的源树。

请在 MSDP 视图进行下列配置。

操作	命令
创建 SA 缓存状态	**cache-sa-enable**
禁止缓存 SA 状态	**undo cache-sa-enable**

在缺省情况下，路由器收到 SA 消息后缓存 SA 的状态。

6.2.6 配置缓存 SA 的最大数量

为防止路由器受到 DoS（Deny of Service）攻击，可以配置路由器缓存 SA 消息的最大数量。

请在 MSDP 视图下进行下列配置。

操作	命令
配置缓存 SA 的最大数量	**peer** *peer-address* **sa-cache-maximum** *sa-limit*
恢复缺省设置	**undo peer** *peer-address* **sa-cache-maximum**

缺省情况下，缓存 SA 的最大值是 2048。

6.2.7 配置请求源消息

当一个新组加入时，路由器向指定的 MSDP 对等体发送 SA 请求消息，该 MSDP 对等体收到 SA 请求消息后，回应其缓存的 SA 信息。

请在 MSDP 视图进行下列配置。

操作	命令
接收到新组加入消息时向指定 MSDP 对等体发送 SA 请求消息	**peer** *peer-address* **request-sa-enable**
恢复缺省设置	**undo peer** *peer-address* **request-sa-enable**

缺省情况下，收到新组的加入消息时，路由器并不立即向其 MSDP 对等体发送 SA 请求消息，而是等待下一个 SA 消息的到来。

6.2.8 配置对源信息的处理

对于源信息，可以从 SA 的创建、转发和对 SA 请求消息的接收等方面进行控制。

1. 配置 SA 通告的组播源

可以配置 MSDP 在创建 SA 消息时，只通告组播路由表中符合条件的（S，G）项，以对引入域内的（S，G）项进行控制。

请在 MSDP 视图进行下列配置。

操作	命令
只通告通过 ACL 的（S，G）项	**import-source** [**acl** *acl-number*]
取消上述配置	**undo import-source**

缺省情况下，SA 消息只通告域内的源。如果执行 **import-source** 命令时不指定 ACL，SA 消息将不通告任何源。

2. 配置接收的 SA 请求消息

通常路由器会接收其 MSDP 对等体发来的所有 SA 请求消息，可以配置路由器是否对所有的 SA 请求都进行响应。

请在 MSDP 视图进行下列配置。

操作	命令
只处理来自指定 MSDP 对等体的通过访问控制列表的 SA request 消息	**peer** *peer-address* **sa-request-policy** [**acl** *acl-number*]
取消过滤 SA 请求消息的配置	**undo peer** *peer-address* **sa-request-policy**

缺省情况下，只有缓存 SA 消息的路由器才能回应 SA 请求消息。如果不指定访问控制列表，则忽略其 MSDP 对等体发送的所有 SA 请求。如果指定了访问控制列表，则只处理符合该列表的组的 SA 请求消息。

3. 通过 MSDP 出方向的过滤器控制 SA 消息转发

请在 MSDP 视图进行下列配置。

操作	命令
配置到指定 MSDP 对等体的 SA 消息是否被转发	**peer** *peer-address* **sa-policy export** [**acl** *acl-number*]
取消对转发的源信息进行过滤	**undo peer** *peer-address* **sa-policy export**

缺省情况下，所有 SA 消息都向 MSDP 对等体转发。如果在执行 **peer sa-policy export** 命令时不指定 **acl**，将不转发任何 SA。如果指定了访问控制列表，则只转发符合该列表的 SA 消息。

4. 通过 TTL 控制 SA 消息转发

对于带封装数据的 SA 消息，只有当它的 IP 头的 TTL 不小于阈值时才能发送到指定的 MSDP 对等体。

请在 MSDP 视图进行下列配置。

操作	命令
配置最小 TTL	**peer** *peer-address* **ttl-threshold** *ttl*
恢复缺省值	**undo peer** *peer-address* **ttl-threshold**

缺省情况下，TTL 阈值为 0。

5. 配置接收的 SA 信息

与 MSDP 出方向的过滤器功能相似，配置 MSDP 入方向的过滤器就可以控制接受的 SA 消息。

请在 MSDP 视图进行下列配置。

操作	命令
配置是否接受收到的 SA 消息	**peer** *peer-address* **sa-policy import** [**acl** *acl-number*]
取消对接收的源信息进行过滤	**undo peer** *peer-address* **sa-policy import**

缺省情况下，从对等体接收到的 SA 消息不进行过滤，全部接收。

6.2.9 配置 MSDP 全连接组

如果需要在几个 MSDP 对等体之间实现完全连接，又要避免 SA 消息的泛滥，就可以将这几个 MSDP 对等体配置为一个全连接组（Mesh Group）。来自组外的 SA 消息发送给组内其他成员，同时，对来自组内对等体的 SA 消息不再进行 Peer-RPF 检查。也不再在组内转发。

请在 MSDP 视图进行下列配置。

操作	命令
配置 MSDP 对等体成为全连接组成员	**peer** *peer-address* **mesh-group** *name*
取消配置的 MSDP 全连接组成员	**undo peer** *peer-address* **mesh-group** *name*

将同一 MSDP 对等体配置到多个全连接组时，最后一个有效。

6.2.10 配置 MSDP 连接重试周期

请在 MSDP 视图下进行下列配置。

操作	命令
配置 MSDP 连接重试周期	**timer retry** *seconds*
恢复 MSDP 连接重试周期的缺省值	**undo timer retry**

缺省情况下，MSDP 连接重试周期为 30 秒。

6.2.11 关闭 MSDP 对等体

可以关闭 MSDP 对等体之间的会话，在需要的时候再进行激活。

关闭 MSDP 对等体之间的会话后，TCP 连接关闭，并不再重试建立连接，但配置信息会保留。

请在 MSDP 视图进行下列配置。

操作	命令
关闭指定的 MSDP 对等体	**shutdown** *peer-address*
取消关闭指定的 MSDP 对等体	**undo shutdown** *peer-address*

缺省情况下，MSDP 对等体是使能的。

6.2.12 复位 MSDP

请在用户视图下进行下列操作。

操作	命令
复位与指定 MSDP 对等体的 TCP 连接	**reset msdp peer** *peer-address*
复位 MSDP 对等体的统计信息	**reset msdp statistics** [*peer-address*]
复位 MSDP 的 SA 缓存项	**reset msdp sa-cache** [*group-address*]

当使用 **reset msdp peer** 命令复位 TCP 连接时，指定 MSDP 对等体的所有统计信息也将同时被复位。

6.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后 MSDP 的运行情况，验证配置的效果。

在用户视图下执行 **debugging** 命令可对 MSDP 进行调试。

操作	命令
显示来自指定自治域的SA 消息中源和组的数量	**display msdp sa-count** [*as-number*]
显示 MSDP 对等体 的详细信息	**display msdp peer-status** [*peer-address*]
显示指定（S，G）项的状态	**display msdp sa-cache** [*group-address* | [*source-address*]] [*autonomous-system-number*]
显示 MSDP 对等体状态	**display msdp brief**
打开 MSDP 的所有调试信息开关	**debugging msdp all**

操作	命令
关闭 MSDP 的所有调试信息开关	**undo debugging msdp all**
打开 MSDP 事件的调试信息开关	**debugging msdp event**
关闭 MSDP 事件的调试信息开关	**undo debugging msdp event**
打开 MSDP 报文的调试信息开关	**debugging msdp packet**
关闭 MSDP 报文的调试信息开关	**undo debugging msdp packet**
打开 MSDP 路由的调试信息开关	**debugging msdp source-active**
关闭 MSDP 路由的调试信息开关	**undo debugging msdp source-active**
打开 MSDP 对等体重建连接的调试信息开关	**debugging msdp connect**
关闭 MSDP 对等体重建连接的调试信息开关	**undo debugging msdp connect**

说明：

只有执行了 **cache-sa-enable-**命令之后，执行 **display msdp sa-count** 命令才会有输出。

6.4 配置实例

6.4.1 配置静态 RPF 对等体

1. 组网需求

在图 6-3所示的环境中，四台路由器均在 PIM-SM 域上，它们之间没有运行 BGP 或 MBGP。

为了使 Router D 能够获得 PIM-SM 域 1、域 2 和域 3 三个域中指定的 SA 消息，可以配置带有 **rp-policy** 参数的静态 RPF 对等体。

配置完成后，Router D 只接收来自其静态 RPF 对等体且被相应过滤策略允许的 SA 消息。

2. 组网图

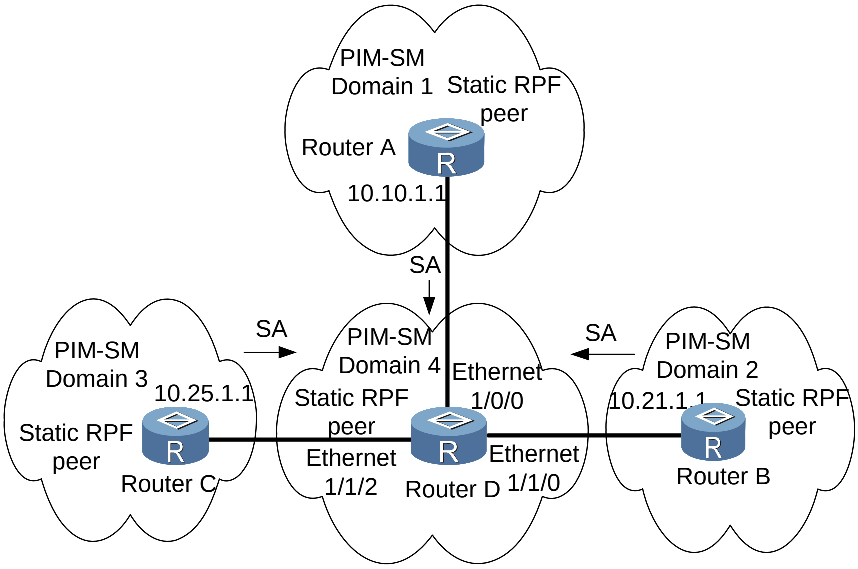


图6-3 配置静态 RPF 对等体组网图

3. 配置步骤

配置路由器 Router A 为 Router D 的静态 RPF 对等体。

```
<Router D> system-view
[Router D] ip ip-prefix list-a permit 10.10.0.0 16
[Router D] msdp
[Router D-msdp] peer 10.10.1.1 connect-interface ethernet 1/0/0
[Router D-msdp] static-rpf-peer 10.10.1.1 rp-policy list-a
```

配置路由器 Router B 为 Router D 的静态 RPF 对等体。

```
[Router D] ip ip-prefix list-b permit 10.21.0.0 16
[Router D] msdp
[Router D-msdp] peer 10.21.1.1 connect-interface ethernet 1/1/0
[Router D-msdp] static-rpf-peer 10.21.1.1 rp-policy list-b
```

配置路由器 Router C 为 Router D 的静态 RPF 对等体。

```
[Router D] ip ip-prefix list-c permit 10.25.0.0 16
[Router D] msdp
[Router D-msdp] peer 10.25.1.1 connect-interface ethernet 1/1/2
[Router D-msdp] static-rpf-peer 10.25.1.1 rp-policy list-c
```

6.4.2 配置 Anycast RP

1. 组网需求

在 PIM-SM 域内配置 Anycast RP。Router A 和 Router B 之间建立 MSDP 对等体，在 Router A 和 Router B 上配置向外发送 SA 消息时使用 Loopback0 地址；设置 Router A 和 Router B 的 Loopback10 接口为 BSR/RP，并配置 Anycast RP 地址，以达到当有组播组成员加入时，与组播组成员直接相连的路由器能够向拓扑距离最近的 RP 发起加入请求的目的。

说明：

本例主要实现 Router A 和 Router B 的配置，由于 Router E、Router D 和 Router C 上的配置主要是启动组播和在各个接口上启动 PIM-SM 协议，本实例将省略此部分内容。

2. 组网图

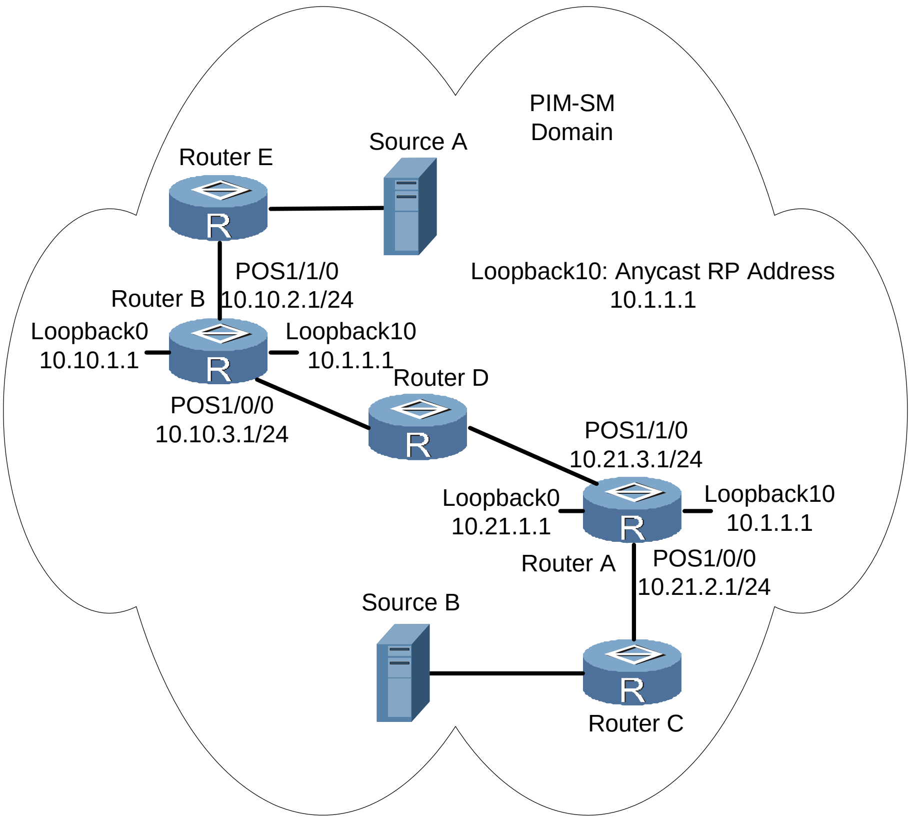


图6-4 配置 Anycast RP 组网图

3. 配置步骤

(1) 配置 Router B

启动组播。

```
<Router B> system-view
[Router B] multicast routing-enable
```

配置接口 Loopback0 的 IP 地址。

```
[Router B] interface loopback 0
[Router B-LoopBack0] ip address 10.10.1.1 255.255.255.255
[Router B-LoopBack0] quit
```

配置接口 Loopback10 的 IP 地址，启动 PIM-SM。

```
[Router B] interface loopback 10
[Router B-LoopBack10] ip address 10.1.1.1 255.255.255.255
[Router B-LoopBack10] pim sm
[Router B-LoopBack10] quit
```

配置接口 POS1/1/0 的 IP 地址，启动 PIM-SM。

```
[Router B] interface pos 1/1/0
[Router B-Pos1/1/0] ip address 10.10.2.1 255.255.255.0
[Router B-Pos1/1/0] pim sm
[Router B-Pos1/1/0] undo shutdown
[Router B-Pos1/1/0] quit
```

配置接口 POS1/0/0 的 IP 地址，启动 PIM-SM。

```
[Router B] interface pos 1/0/0
[Router B-Pos1/0/0] ip address 10.10.3.1 255.255.255.0
[Router B-Pos1/0/0] pim sm
[Router B-Pos1/0/0] undo shutdown
[Router B-Pos1/0/0] quit
```

配置 OSPF。

```
[Router B] ospf
[Router B-ospf] area 0
[Router B-ospf-area-0.0.0.0] network 10.10.2.0 0.255.255.255
[Router B-ospf-area-0.0.0.0] network 10.10.3.0 0.255.255.255
[Router B-ospf-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[Router B-ospf-area-0.0.0.0] network 10.10.1.1 0.0.0.0
[Router B-ospf-area-0.0.0.0] quit
[Router B-ospf] quit
```

配置 Router A 为 Router B 的 MSDP 对等体。

```
[Router B] msdp
[Router B-msdp] peer 10.21.1.1 connect-interface loopback 0
```

配置 Originating RP。

```
[Router B-msdp] originating-rp loopback 0
```

```
[Router B-msdp] quit
```

配置候选 RP 和 BSR。

```
[Router B] pim
[Router B-pim] c-rp loopback 10
[Router B-pim] c-bsr loopback 10 30
```

(2) 配置 Router A

启动组播。

```
[Router A] multicast routing-enable
```

配置接口 Loopback0 的 IP 地址。

```
[Router A] interface loopback 0
[Router A-LoopBack0] ip address 10.21.1.1 255.255.255.255
[Router A-LoopBack0] quit
```

配置接口 Loopback10 的 IP 地址，启动 PIM-SM。

```
[Router A] interface loopback10
[Router A-LoopBack10] ip address 10.1.1.1 255.255.255.255
[Router A-LoopBack10] pim sm
[Router A-LoopBack10] quit
```

配置接口 POS1/0/0 的 IP 地址，启动 PIM-SM。

```
[Router A] interface pos 1/0/0
[Router A-Pos1/0/0] ip address 10.21.2.1 255.255.255.0
[Router A-Pos1/0/0] pim sm
[Router A-Pos1/0/0] undo shutdown
[Router A-Pos1/0/0] quit
```

配置接口 POS1/1/0 的 IP 地址，启动 PIM-SM。

```
[Router A] interface pos 1/1/0
[Router A-Pos1/1/0] ip address 10.21.3.1 255.255.255.0
[Router A-Pos1/1/0] pim sm
[Router A-Pos1/1/0] undo shutdown
[Router A-Pos1/1/0] quit
```

配置 OSPF 路由。

```
[Router A] ospf
[Router A-ospf] area 0
[Router A-ospf-area-0.0.0.0] network 10.21.2.0 0.255.255.255
[Router A-ospf-area-0.0.0.0] network 10.21.3.0 0.255.255.255
[Router A-ospf-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[Router A-ospf-area-0.0.0.0] network 10.21.1.1 0.0.0.0
[Router A-ospf-area-0.0.0.0] quit
[Router A-ospf] quit
```

配置 Router B 为 Router A 的 MSDP 对等体。

```
[Router A] msdp
[Router A-msdp] peer 10.10.1.1 connect-interface loopback 0
```

配置 Originating RP。

```
[Router A-msdp] originating-rp loopback 0
[Router A-msdp] quit
```

配置候选 RP 和 BSR。

```
[Router A] pim
[Router A-pim] c-rp loopback 10
[Router A-pim] c-bsr loopback 10 30
```

6.4.3 MSDP 综合组网

1. 组网需求

在如图 6-5网络中，运行 MSDP，在 PIM-SM 域 1 内配置 Anycast RP，PIM-SM 各个域的 RP 之间建立 MSDP 对等体关系，域间使用 MBGP 组播扩展路由协议（此部分命令请参见“第 7 章 配置 MBGP 组播扩展”）。

2. 组网图

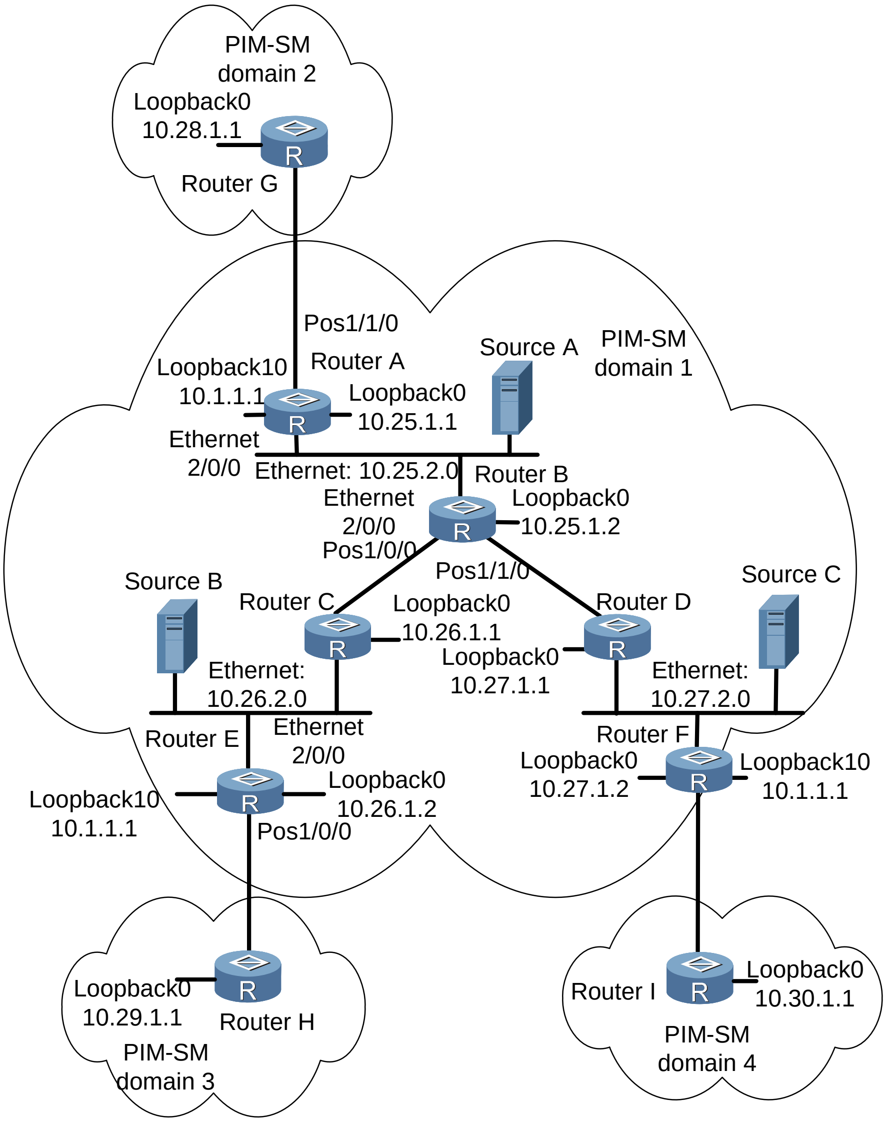


图6-5 MSDP 综合组网案例组网图

3. 配置步骤

(1) 配置 Router A

启动组播。

```
<Router A> system-view
[Router A] multicast routing-enable
```

配置接口 loopback0 的 IP 地址，启动 PIM-SM。

```
[Router A] interface loopback0
[Router A-LoopBack0] ip address 10.25.1.1 255.255.255.255
[Router A-LoopBack0] pim sm
[Router A-LoopBack0] quit
```

配置接口 loopback10 的 IP 地址，启动 PIM-SM。

```
[Router A] interface loopback10
[Router A-LoopBack10] ip address 10.1.1.1 255.255.255.255
[Router A-LoopBack10] pim sm
[Router A-LoopBack10] quit
```

配置接口 ethernet 2/0/0 的 IP 地址，启动 PIM-SM。

```
[Router A] interface ethernet 2/0/0
[Router A-Ethernet2/0/0] ip address 10.25.2.3 255.255.255.0
[Router A-Ethernet2/0/0] pim sm
[Router A-Ethernet2/0/0] undo shutdown
[Router A-Ethernet2/0/0] quit
```

配置接口 pos 1/1/0 的 IP 地址，启动 PIM-SM。

```
[Router A] interface pos 1/1/0
[Router A-Pos1/1/0] ip address 10.25.3.1 255.255.255.0
[Router A-Pos1/1/0] pim sm
[Router A-Pos1/1/0] undo shutdown
[Router A-Pos1/1/0] quit
```

配置 OSPF。

```
[Router A] ospf
[Router A-ospf] area 0
[Router A-ospf-area-0.0.0.0] network 10.25.2.0 0.255.255.255
[Router A-ospf-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[Router A-ospf-area-0.0.0.0] network 10.25.1.1 0.0.0.0
[Router A-ospf-area-0.0.0.0] quit
[Router A-ospf] quit
```

配置 BGP。

```
[Router A] bgp 100
[Router A-bgp] group rtagroup26 external
[Router A-bgp] peer 10.26.1.2 group rtagroup26 as-number 1
[Router A-bgp] peer rtagroup26 connect-interface loopback0
[Router A-bgp] ipv4-family multicast
[Router A-bgp-af-mul] peer rtagroup26 enable
[Router A-bgp-af-mul] peer rtagroup26 next-hop-local
[Router A-bgp-af-mul] quit
[Router A-bgp] peer rtagroup26 enable
[Router A-bgp] peer rtagroup26 next-hop-local
[Router A-bgp] peer rtagroup26 ebgp-max-hop 255
[Router A-bgp] group rtagroup27 external
[Router A-bgp] peer 10.27.1.2 group rtagroup27 as-number 1
[Router A-bgp] peer rtagroup27 connect-interface loopback0
```

```
[Router A-bgp] ipv4-family multicast
[Router A-bgp-af-mul] peer rtagroup27 enable
[Router A-bgp-af-mul] peer rtagroup27 next-hop-local
[Router A-bgp-af-mul] quit
[Router A-bgp] peer rtagroup27 enable
[Router A-bgp] peer rtagroup27 next-hop-local
[Router A-bgp] peer rtagroup27 ebgp-max-hop 255
[Router A-bgp] group rtagroup28 external
[Router A-bgp] peer 10.28.1.1 group rtagroup28 as-number 2
[Router A-bgp] ipv4-family multicast
[Router A-bgp-af-mul] peer rtagroup28 enable
[Router A-bgp-af-mul] peer rtagroup28 next-hop-local
[Router A-bgp-af-mul] quit
[Router A-bgp] peer rtagroup28 enable
[Router A-bgp] peer rtagroup28 next-hop-local
[Router A-bgp] peer rtagroup28 default-route-advertise
[Router A-bgp] peer rtagroup28 ebgp-max-hop 255
[Router A-bgp] quit
```

配置 MSDP 对等体、全连接组和 Originating RP。

```
[Router A] msdp
[Router A-msdp] peer 10.28.1.1 connect-interface loopback 0
[Router A-msdp] peer 10.26.1.2 connect-interface loopback 0
[Router A-msdp] peer 10.27.1.2 connect-interface loopback 0
[Router A-msdp] peer 10.26.1.2 mesh-group net
[Router A-msdp] peer 10.27.1.2 mesh-group net
[Router A-msdp] originating-rp loopback0
[Router A-msdp] quit
```

配置候选 RP 和 BSR。

```
[Router A] pim
[Router A-pim] c-rp loopback 10
[Router A-pim] c-bsr loopback 0 30
```

(2) 配置 Router E

启动组播。

```
[Router E] multicast routing-enable
```

配置接口 IP 地址，启动 PIM-SM。

```
[Router E] interface loopback0
[Router E-LoopBack0] ip address 10.26.0.1.2 255.255.255.255
[Router E-LoopBack0] pim sm
[Router E-LoopBack0] quit
```

配置接口 IP 地址，启动 PIM-SM。

```
[Router E] interface loopback10
[Router E-LoopBack10] ip address 10.1.1.1 255.255.255.255
[Router E-LoopBack10] pim sm
[Router E-LoopBack10] quit
```

配置接口 IP 地址，启动 PIM-SM。

```
[Router E] interface ethernet 2/0/0
[Router E-Ethernet2/0/0] ip address 10.26.2.3 255.255.255.0
[Router E-Ethernet2/0/0] pim sm
[Router E-Ethernet2/0/0] undo shutdown
[Router E-Ethernet2/0/0] quit
```

配置接口 IP 地址，启动 PIM-SM。

```
[Router E] interface pos 1/0/0
[Router E-Pos1/0/0] ip address 10.26.3.1 255.255.255.0
[Router E-Pos1/0/0] pim sm
[Router E-Pos1/0/0] undo shutdown
[Router E-Pos1/0/0] quit
```

配置 OSPF。

```
[Router E] ospf
[Router E-ospf] area 0
[Router E-ospf-area-0.0.0.0] network 10.26.2.0 0.255.255.255
[Router E-ospf-area-0.0.0.0] network 10.1.1.1 0.0.0.0
[Router E-ospf-area-0.0.0.0] network 10.26.1.2 0.0.0.0
 [Router E-ospf-area-0.0.0.0] quit
[Router E-ospf] quit
```

配置 BGP。

```
[Router E] bgp 100
[Router E-bgp] group rtegroup25 external
[Router E-bgp] peer 10.25.1.1 group rtegroup25 as-number 1
[Router E-bgp] peer rtegroup25 connect-interface loopback0
[Router E-bgp] ipv4-family multicast
[Router E-bgp-af-mul] peer rtegroup25 enable
[Router E-bgp-af-mul] peer rtegroup25 next-hop-local
[Router E-bgp-af-mul] quit
[Router E-bgp] peer rtegroup25 enable
[Router E-bgp] peer rtegroup25 next-hop-local
[Router E-bgp] peer rtegroup25 ebgp-max-hop 255
[Router E-bgp] group rtegroup27 external
[Router E-bgp] peer 10.27.1.2 group rtegroup27 as-number 1
[Router E-bgp] peer rtegroup27 connect-interface loopback0
```

```
[Router E-bgp] ipv4-family multicast
[Router E-bgp-af-mul] peer rtegroup27 enable
[Router E-bgp-af-mul] peer rtegroup27 next-hop-local
[Router E-bgp-af-mul] quit
[Router E-bgp] peer rtegroup27 enable
[Router E-bgp] peer rtegroup27 next-hop-local
[Router E-bgp] peer rtegroup27 ebgp-max-hop 255
[Router E-bgp] quit
```

配置 MSDP 对等体、全连接组和 Originating RP。

```
[Router E] msdp
[Router E-msdp] peer 10.29.1.1 connect-interface loopback 0
[Router E-msdp] static-rpf-peer 10.29.1.1
[Router E-msdp] peer 10.25.1.1 connect-interface loopback 0
[Router E-msdp] peer 10.27.1.2 connect- interface loopback 0
[Router E-msdp] peer 10.25.1.1 mesh-group net
[Router E-msdp] peer 10.27.1.2 mesh-group net
[Router E-msdp] originating-rp loopback0
[Router E-msdp] quit
[Router E] ip route-static 10.29.1.1 255.255.255.0 pos1/0/0
```

配置候选 RP 和 BSR。

```
[Router E] pim
[Router E-pim] c-rp loopback 10
[Router E-pim] c-bsr loopback 0 30
```

(3) 配置 Router B

启动组播。

```
[Router B] multicast routing-enable
```

配置接口 IP 地址，启动 PIM-SM。

```
[Router B] interface ethernet 2/0/0
[Router B-Ethernet2/0/0] ip address 10.25.2.4 255.255.255.0
[Router B-Ethernet2/0/0] pim sm
[Router B-Ethernet2/0/0] undo shutdown
```

第7章 配置 MBGP 组播扩展

本章包括以下内容：

- MBGP 组播扩展简介
- 配置内容
- 显示和调试
- 配置实例

7.1 MBGP 组播扩展简介

7.1.1 MBGP 组播扩展概述

目前使用最多的域间单播路由协议是 BGP-4，由于组播的网络拓扑和单播拓扑有可能不同，为了实现域间组播路由信息的传递，必须对 BGP-4 进行改动。网络中的一些路由器可能只支持单播不支持组播，也可能按照策略配置不转发组播分组。为了构造域间组播路由树，除了要知道单播路由信息外，还要知道网络中哪些部分是支持组播的，即组播的网络拓扑情况。

BGP-4 已被证明是一个有效、稳定的单播域间路由协议，因此合理的解决方案是对 BGP-4 协议进行增强和扩展，而不是构建一套全新的协议。RFC2858 中规定了对 BGP 进行多协议扩展的方法，扩展后的 BGP 被称为 MBGP（Multi-protocol BGP）或者 BGP-4+。

MBGP 不仅能携带 IPv4 单播路由信息，也能携带其他网络层协议（如组播、IPv6 等）的路由信息，携带组播路由信息只是其中一个扩展功能。本章主要介绍 MBGP 组播扩展。

有了 MBGP 之后，单播和组播路由信息可以通过同一个进程交换，存放在不同的路由表里。由于 MBGP 是 BGP-4 协议的一个增强版，因此 BGP-4 所支持的常见的策略和配置方法都可以用到组播里。

7.1.2 MBGP 为组播扩展的属性

为了在 MBGP 中支持组播能力，RFC2858 在 Update 报文里定义了两个新的路径属性：MP_REACH_NLRI（Multi-protocol Reachable NLRI）和 MP_UNREACH_NLRI（Multi-protocol Unreachable NLRI）。

这两个属性都是可选非过渡（optional non-transitive）属性，即不支持 MBGP 的路由器可以忽略这些属性里的信息，不转发这些属性。

在 MP_REACH_NLRI 和 MP_UNREACH_NLRI 属性中所携带的信息中，包含了地址族标识 AFI（Address Family Identifier）和补充地址族标识 SAFI（Subsequent Address Family Identifier）字段，可以指定属性信息是针对哪个地址族的。

SAFI 是对 NLRI（Network Layer Reachability Information）中信息的补充，在 RFC2858 中定义了三个值：

- 1 表示 NLRI 中的信息是单播模式；
- 2 表示 NLRI 中的信息是组播模式；
- 3 则表示 NLRI 域中的信息即可用于单播也可用于组播。

1. MP_REACH_NLRI 属性

MP_REACH_NLRI 是可选非过渡属性，具有以下用途：

- 发送可达新协议的路由；
- 发送新协议的下一跳信息，编码方式与 MP_NLRI 相同；
- 允许路由器报告部分或全部存在于本地系统的子网接入点 SNPA（Sub-network Points of Attachment）；

包含 MP_REACH_NLRI 属性的 Update 报文必须包含 ORIGIN 属性和 AS_PATH 属性。

2. MP_UNREACH_NLRI 属性

MP_UNREACH_NLRI 属性是可选非过渡属性，用于取消一条或多条不可达路由，包含的字段有：

- AFI、SAFI；
- Withdrawn Routes：取消的路由，其中包括一个或多个 NLRI，NLRI 里的内容是不可达的目的地址。

包含 MP_UNREACH_NLRI 属性的 UPDATE 报文不要求携带其他路径属性。

由于增加了这两个属性，MBGP 具有携带多协议信息的功能，既可以支持单播又可以支持组播，在网络中构建不同的拓扑结构，而且能分别支持不同的策略。所以，针对某一个策略，MBGP 构造出来的域间单播和组播转发路由可能是不同的。

7.1.3 MBGP 的运行方式与消息类型

MBGP 在路由器上以 IBGP（Internal BGP）和 EBGP（External BGP）两种方式运行。当 MBGP 运行于同一自治系统（AS）内部时，被称为 IBGP；当 MBGP 运行于不同自治系统之间时，称为 EBGP。

MBGP 有四种消息：

- Open Message：TCP 连接建立后第一个发送的消息，它用于建立 BGP 对等体间的连接关系。
- Update Message：MBGP 最重要的信息，用于在对等体之间交换路由信息，它最多由三部分构成：不可达路由（MP_UNREACH_NLRI）、路径属性（Path Attributes）、网络可达性信息（MP_REACH_NLRI）。
- Notification Message：错误通告消息。
- Keepalive Message：用于检测连接有效性。

7.2 配置内容

配置 MBGP 组播扩展包括以下内容：

- 启动 MBGP 组播扩展协议
- 指定 MBGP 组播扩展要通告的网络路由
- 配置 MBGP 对等体/对等体组
- 配置 MBGP 路由聚合
- 配置 MBGP 引入 IGP 的路由信息

下面是一些使用方法与 BGP 基本相同的配置，具体内容请参考“路由协议配置”部分的“第 6 章 配置 BGP”和“第 8 章 配置 IP 路由策略”。

- 配置自治系统的 MED 值
- 比较来自不同自治系统邻居路径的 MED 值
- 配置本地优先级
- 配置 MBGP 定时器
- 配置 MBGP 路由反射器
- 配置 MBGP 团体属性
- 定义 AS 路径列表和路由策略
- 配置 MBGP 路由过滤
- 复位 BGP 连接

7.2.1 启动 MBGP 组播扩展协议

启动 MBGP 组播扩展协议时应进入 IPv4 地址族的组播子地址族。

启动 MBGP 组播扩展协议后，本地路由器不接收相邻路由器的 MBGP 连接请求。去掉 MBGP 组播地址族时，MBGP 将切断所有已经建立的 MBGP 连接。

请在 BGP 视图下进行下列配置。

操作	命令
进入 MBGP 组播地址族模式	**ipv4-family multicast**
去掉 MBGP 组播地址族模式	**undo ipv4-family multicast**

缺省情况下，系统不运行 MBGP 组播扩展协议。

7.2.2 指定 MBGP 组播扩展要通告的网络路由

用 **network** 命令来指定 MBGP 要向对等体通告的网络路由，同时也可指定此网络路由的掩码和使用的 Route-policy。

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置本地 MBGP 要通告的网络路由	**network** *ip-address* [*address-mask*] [**route-policy** *policy-name*]
取消本地 MBGP 要通告的网络路由	**undo network** *ip-address* [*address-mask*] [**route-policy** *policy-name*]

缺省情况下，本地 MBGP 不通告任何路由。

使用 **network** 命令通告的是精确匹配的路由，即前缀与掩码要与配置完全符合才能正确发布。如果没有指定掩码，则按照自然网段进行精确匹配。

7.2.3 配置 MBGP 对等体/对等体组

MBGP 对等体组的使用是为了方便用户配置。当用户启动若干配置相同的对等体时，可先创建一对等体组并将其配置好；然后将各对等体组加入到此对等体组中以使其获得与此对等体组相同的配置。

注意：

请在技术支持人员的指导下配置对等体组。

1. 创建对等体组及加入组员

缺省情况下，IBGP 对等体会加入到一个默认的对等体组中，不需要进行配置。这个默认的对等体组是不可见的，对任何一个 IBGP 对等体的路由更新策略的配置只对其所在组内的其他 IBGP 对等体有效。若路由器不是路由反射器，所有 IBGP 对等体在同一个组内；否则，所有的路由反射客户在一个组内，非客户在另一个组内。

EBGP 对等体组的成员必须在同一网段内，否则某些 EBGP 对等体可能会丢弃本路由器发送的路由更新报文。

不能将 IBGP 对等体和 EBGP 对等体加入同一个组中。

请在 BGP 视图下进行下列配置。

操作	命令
创建一个对等体组	**group** *group-name* [**external** | **internal**]
删除指定的对等体组	**undo group** *group-name*
在对等体组中创建一个对等体	**peer** *peer-address* **group** *group-name*
从对等体组中删除一个对等体	**undo peer** *peer-address* **group** *group-name*
复位对等体组所有成员的连接(用户视图)	**reset bgp group** *group-name*

如果设备上创建的对等体组是 **external**，则对等体组在该设备外部；若创建的对等体组是 **internal**，则对等体组包含该设备在内。

2. 使能对等体/对等体组

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
使能指定对等体/对等体组	**peer** { *peer-address* | *group-name* } **enable**
去使能指定对等体/对等体组	**undo peer** { *peer-address* | *group-name* } **enable**

3. 配置将 MBGP 团体属性传给对等体/对等体组

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置将团体属性传给对等体/对等体组	**peer** { *group-name* | *peer-address* } **advertise-community**
不将团体属性传给对等体/对等体组	**undo peer** { *group-name* | *peer-address* } **advertise-community**

缺省情况下，不将团体属性传给任何对等体/对等体组。

4. 配置对等体/对等体组为 MBGP 路由反射器的客户

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置对等体/对等体组作为 MBGP 路由反射器客户	**peer** { *group-name* | *peer-address* } **reflect-client**
取消对等体/对等体组作为 MBGP 路由反射器客户	**undo peer** { *group-name* | *peer-address* } **reflect-client**

缺省情况下，自治系统中无路由反射器。

一般来说，不需要为对等体组配置此命令。保留该项配置是为了在某些情况下与其他厂商的设备兼容。

5. 配置在发布路由时将自身地址作为下一跳

取消向对等体/对等体组发布的路由中下一跳的处理，把自身地址作为下一跳。此配置只对 IBGP 对等体/对等体组有效。

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
发布路由时将自身地址作为下一跳	**peer** { *peer-address* | *group-name* } **next-hop-local**
发布路由时不将自身地址作为下一跳	**undo peer** { *peer-address* | *group-name* } **next-hop-local**

6. 指定对等体/对等体组的路由策略

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
指定对等体/对等体组的路由策略	**peer** { *group-name* | *peer-address* } **route-policy** *policy-name* { **import** | **export** }
删除指定对等体/对等体组的路由策略	**undo peer** { *group-name* | *peer-address* } **route-policy** *policy-name* { **import** | **export** }

缺省情况下，不指定对等体/对等体组的路由策略。

7. 配置对等体/对等体组的基于 IP 访问控制列表的路由过滤策略

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置对等体/对等体组的基于 IP 访问控制列表的路由过滤策略	**peer** { *group-name* | *peer-address* } **filter-policy** *acl-number* { **import** | **export** }
删除对等体/对等体组的基于 IP 访问控制列表的路由过滤策略	**undo peer** { *group-name* | *peer-address* } **filter-policy** *acl-number* { **import** | **export** }

缺省情况下，对等体/对等体组不基于 IP 访问控制列表进行路由过滤。

8. 配置对等体/对等体组的基于 AS 路径列表的路由过滤策略

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置对等体/对等体组的基于 AS 访问列表的过滤策略	**peer** { *group-name* | *peer-address* } **as-path-acl** *number* { **import** | **export** }
删除对等体/对等体组的基于 AS 访问列表的过滤策略	**undo peer** { *group-name* | *peer-address* } **as-path-acl** *number* { **import** | **export** }

缺省情况下，对等体/对等体组不基于 AS 访问列表进行路由过滤。

9. 配置对等体/对等体组的基于前缀列表的路由过滤策略

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置对等体/对等体组的基于前缀列表的过滤策略	**peer** { *group-name* | *peer-address* } **ip-prefix** *prefixname* { **import** | **export** }
删除对等体/对等体组的基于前缀列表的过滤策略	**undo peer** { *group-name* | *peer-address* } **ip-prefix** *prefixname* { **import** | **export** }

缺省情况下，对等体/对等体组不基于前缀列表进行路由过滤。

7.2.4 配置 MBGP 路由聚合

MBGP 支持路由的手动聚合。手动聚合是对 MBGP 本地路由进行的聚合，在手动聚合方式下还可进行一系列参数的配置。

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置本地路由聚合	**aggregate** *address mask* [**as-set**] [**detail-suppressed**] [**suppress-policy** *route-policy-name*] [**origin-policy** *route-policy-name*] [**attribute-policy** *route-policy-name*]
取消本地路由聚合	**undo aggregate** *address mask* [**as-set**] [**detail-suppressed**] [**suppress-policy** *route-policy-name*] [**origin-policy** *route-policy-name*] [**attribute-policy** *route-policy-name*]

缺省情况下，MBGP 不对本地路由进行聚合。

7.2.5 配置 MBGP 引入 IGP 的路由信息

MBGP 可以向其他的自治区域发送本区域的内部网络的信息。为了达到此目的，可以将本地路由器通过 IGP 路由协议得到的关于本区域内部的网络信息通过 MBGP 发送出去。

请在 IPv4 组播子地址族视图下进行下列配置。

操作	命令
配置 MBGP 引入 IGP 的路由信息	**import-route** *protocol* [**route-policy** *policy-name*] [**med** *metric*]
取消引入 IGP 的路由信息	**undo import-route** *protocol*

缺省情况下，MBGP 不引入其他协议的路由。

protocol 指定可引入的源路由协议，可以是 direct、static、rip、isis、ospf 与 ospf-ase、ospf-nssa。

7.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后 MBGP 的运行情况，验证配置的效果。

在用户视图下，执行 **debugging** 命令可对 MBGP 进行调试。

操作	命令
显示 MBGP 路由表中的信息	**display bgp multicast routing-table** *ip-address*
显示 CIDR 路由	**display bgp multicast routing-table cidr**
显示指定 MBGP 团体的路由信息	**display bgp multicast routing-table community** [*community-number* | **no-export-subconfed** | **no-advertise** | **no-export** | **whole-match**]
显示指定 MBGP 团体列表允许的路由信息	**display bgp multicast routing-table community-list** *community-list-number* [**whole-match**]
显示具有不一致的源 AS 的路由	**display bgp multicast routing-table different-origin-as**
显示对等体信息	**display bgp multicast peer** [*peer-address*] **verbose**
显示通过 MBGP 通告的路由信息	**display bgp multicast network**
显示对等体组信息	**display bgp multicast group** [*group-name*]
显示 MBGP 综合信息	**display bgp multicast peer**
打开 MBGP UPDATE 报文调试信息开关	**debugging bgp mp-update**
关闭 MBGP UPDATE 报文调试信息开关	**undo debugging bgp mp-update**

7.4 配置实例

1. 组网需求

本例说明如何通过 MBGP 属性来管理路由选择。

所有路由器都配置 MBGP。AS200 中的 IGP 使用 OSPF。路由器 A 在 AS100 中，并作为 AS200 中的路由器 B 和路由器 C 的 MBGP 邻居。路由器 B 和路由器 C 对路由器 D 运行 IBGP，路由器 D 也在 AS200 中。

2. 组网图

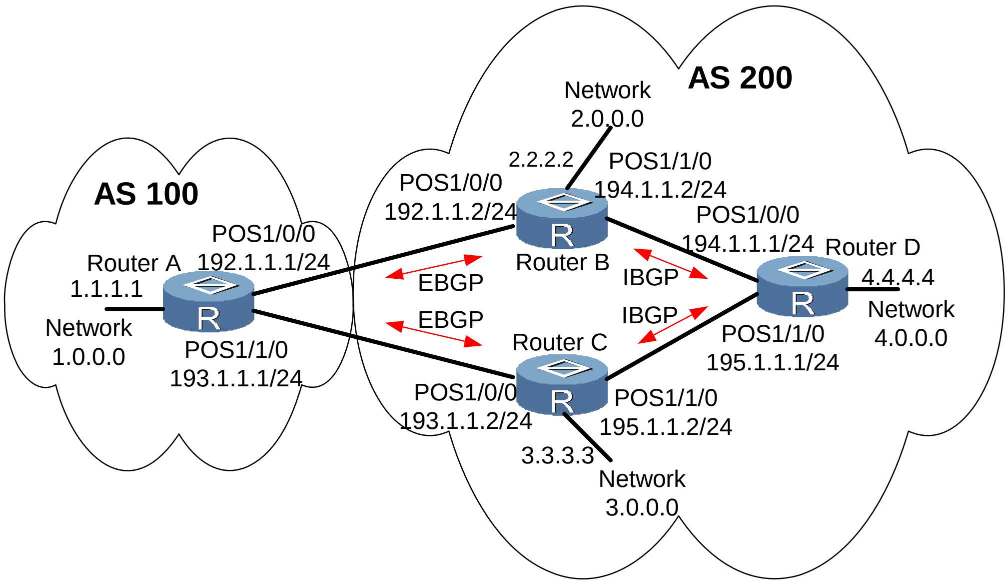


图7-1 配置 MBGP 路径选择的组网图

3. 配置步骤

(1) 配置 Router A

```
<Router A> system-view
[Router A] interface pos 1/0/0
[Router A-Pos1/0/0] ip address 192.1.1.1 255.255.255.0
[Router A-Pos1/0/0] quit
[Router A] interface pos 1/1/0
[Router A-Pos1/1/0] ip address 193.1.1.1 255.255.255.0
[Router A-Pos1/1/0] quit
```

启动 MBGP。

```
[Router A] bgp 100
[Router A-bgp] ipv4-family multicast
```

指定 MBGP 要发送的网络。

```
[Router A-bgp-af-mul] network 1.0.0.0
[Router A-bgp-af-mul] network 2.0.0.0
[Router A-bgp-af-mul] quit
```

配置对等体。

```
[Router A-bgp] group as200 external
[Router A-bgp] peer as200 as-number 200
[Router A-bgp] peer 192.1.1.2 group as200
[Router A-bgp] peer 193.1.1.2 group as200
[Router A-bgp] ipv4-family multicast
[Router A-bgp-af-mul] peer as200 enable
[Router A-bgp-af-mul] peer as200 next-hop-local
```

配置路由器 A 的 MED 属性。

路由器 A 上增加访问控制列表，允许网络 1.0.0.0。

```
[Router A] acl number 2000
[Router A-acl-basic-2000] rule permit ip source1.0.0.0 0.255.255.255
```

定义两个路由策略，一个名为 set_med_50，另一个名为 set_med_100，第一个路由策略为网络 1.0.0.0 设置的 MED 属性为 50，第二个的 MED 属性为 100。

```
[Router A] route-policy set_med_50 permit node 10
[Router A-route-policy] if-match acl 2000
[Router A-route-policy] apply cost 50
[Router A-route-policy] quit
[Router A] route-policy set_med_100 permit node 10
[Router A-route-policy] if-match acl 2000
[Router A-route-policy] apply cost 100
```

应用路由策略 set_med_50 到路由器 C（193.1.1.2）出口路由更新上，应用路由策略 set_med_100 到路由器 B（192.1.1.2）的出口路由更新上。

```
[Router A] bgp 100
[Router A-bgp] ipv4-family multicast
[Router A-bgp-af-mul] peer 193.1.1.2 route-policy set_med_50 export
[Router A-bgp-af-mul] peer 192.1.1.2 route-policy set_med_100 export
```

(2) 配置 Router B

```
<Router B> system-view
[Router B] interface pos 1/0/0
[Router B-Pos1/0/0] ip address 192.1.1.2 255.255.255.0
[Router B-Pos1/0/0] quit
[Router B] interface pos 1/1/0
[Router B-Pos1/1/0] ip address 194.1.1.2 255.255.255.0
[Router B-Pos1/1/0] quit
```

```
[Router B] ospf
[Router B-ospf] area 0
[Router B-ospf-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[Router B-ospf-area-0.0.0.0] network 192.1.1.0 0.0.0.255
[Router B-ospf-area-0.0.0.0] quit
[Router B-ospf] quit
[Router B] bgp 200
[Router B-bgp] group as100 external
[Router B-bgp] peer as100 as-number 100
[Router B-bgp] peer 192.1.1.1 group as100
[Router B-bgp] group as200 internal
[Router B-bgp] peer 194.1.1.1 group as200
[Router B-bgp] peer 195.1.1.2 group as200
[Router B-bgp] ipv4-family multicast
[Router B-bgp-af-mul] peer as100 enable
[Router B-bgp-af-mul] peer as200 enable
```

(3) 配置 Router C

```
<Router C> system-view
[Router C] interface pos 1/0/0
[Router C-Pos1/0/0] ip address 193.1.1.2 255.255.255.0
[Router C-Pos1/0/0] quit
[Router C] interface pos 1/1/0
[Router C-Pos1/1/0] ip address 195.1.1.2 255.255.255.0
[Router C-Pos1/1/0] quit
[Router C] ospf
[Router C-ospf] area 0
[Router C-ospf-area-0.0.0.0] network 193.1.1.0 0.0.0.255
[Router C-ospf-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[Router C-ospf-area-0.0.0.0] quit
[Router C-ospf] quit
[Router C] bgp 200
[Router C-bgp] group as100 external
[Router C-bgp] peer as100 as-number 100
[Router C-bgp] peer 193.1.1.1 group as100
[Router C-bgp] group as200 internal
[Router C-bgp] peer 194.1.1.2 group as200
[Router C-bgp] peer 195.1.1.1 group as200
[Router C-bgp] ipv4-family multicast
[Router C-bgp-af-mul] peer as100 enable
[Router C-bgp-af-mul] peer as200 enable
```

配置路由器 C 的本地优先级属性。

在路由器 C 上加上访问列表 2000，允许网络 1.0.0.0。

```
[Router C] acl number 2000
[Router C-acl-basic-2000] rule permit source 1.0.0.0 0.255.255.255
[Router C-acl-basic-2000] quit
```

定义名为 localpref 的路由策略，该策略中设置了匹配 ACL 2000（访问控制列表 2000）的路由的本地优先级为 200，不匹配的为 100。

```
[Router C] route-policy localpref permit node 10
[Router C-route-policy] if-match acl 2000
[Router C-route-policy] apply local-preference 200
[Router C-route-policy] quit
[Router C] route-policy localpref permit node 20
[Router C-route-policy] apply local-preference 100
```

应用此路由策略到来自 BGP 邻居 193.1.1.2（路由器 A）上的入口流量上。

```
[Router C] bgp 200
[Router C-bgp] ipv4-family multicast
[Router C-bgp-af-mul] peer 193.1.1.1 route-policy localpref import
```

(4) 配置 Router D

```
<Router D> system-view
[Router D] interface pos 1/0/0
[Router D-Pos1/0/0] ip address 194.1.1.1 255.255.255.0
[Router D-Pos1/0/0] quit
[Router D] interface pos 1/1/0
[Router D-Pos1/1/0] ip address 195.1.1.1 255.255.255.0
[Router D-Pos1/1/0] quit
[Router D] ospf
[Router D-ospf] area 0
[Router D-ospf-area-0.0.0.0] network 194.1.1.0 0.0.0.255
[Router D-ospf-area-0.0.0.0] network 195.1.1.0 0.0.0.255
[Router D-ospf-area-0.0.0.0] network 4.0.0.0 0.0.0.255
[Router D-ospf-area-0.0.0.0] quit
[Router D-ospf] quit
[Router D] bgp 200
[Router D-bgp] group as100 external
[Router D-bgp] peer as100 as-number 100
[Router D-bgp] peer 194.1.1.2 group as100
[Router D-bgp] group as200 internal
[Router D-bgp] peer 195.1.1.2 group as200
```

```
[Router D-bgp] ipv4-family multicast
[Router D-bgp-af-mul] peer as100 enable
[Router D-bgp-af-mul] peer as200 enable
```

为使配置生效，所有的 MBGP 邻居需要使用 **reset bgp all** 命令重设。

第8章 配置组播静态路由

本章包括以下内容：

- 组播静态路由简介
- 配置内容
- 显示和调试
- 配置实例

8.1 组播静态路由简介

组播的网络拓扑和单播拓扑有可能不同，网络中的一些路由器可能只支持单播不支持组播。如果需要单播数据包和组播数据包发送的路径不同，则可以配置组播静态路由，将组播路径与单播路径区别开来。

例如当在源和目的端之间的路径上存在不支持组播的路由器时，组播路由器可以使用隧道方式将组播包传送给相邻的组播路由器。

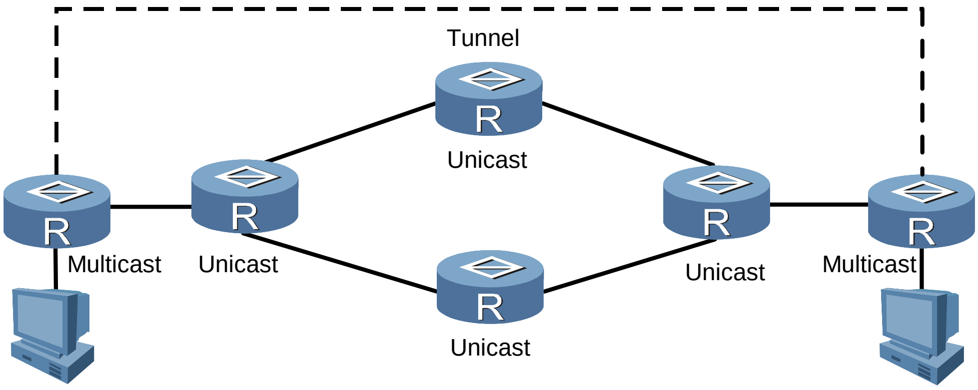


图8-1 隧道方式发送组播数据包

如图 8-1所示，单播路由器（Unicast）仅支持单播，组播数据包由组播路由器（Multicast）通过隧道来转发。如果不配置组播静态路由，则源端（目的端）发送的单播数据包也将通过该隧道传输。组播静态路由解决了这个问题，在组播路由器上配置了一条组播静态路由后，可以实现仅使用隧道发送组播数据包，而对于单播数据包，可以不使用此隧道发送。

组播静态路由仅对所配置的组播路由器生效，不会以任何方式被广播或者引入给其他路由器。

8.2 配置内容

配置组播静态路由包括以下内容：

- 配置组播静态路由
- 改变组播 RPF 路由选择策略

📖 **说明：**

对于同一网段的多条配置，匹配时是按照配置的先后顺序进行的，保证正确的配置顺序对组播是非常重要的。

8.2.1 配置组播静态路由

请在系统视图下进行下列配置。

操作	命令
配置组播静态路由	**ip rpf-route-static** *source* { *mask* | *mask-length* } { *protocol* | **route-policy** *policyname* | *rpf-address* | *interface-type interface-number* } [**order** *order-num* | **preference** *preference*]
删除组播静态路由表中存在的组播静态路由	**undo ip rpf-route-static** *source* { *mask* | *mask-length* } [*protocol*] [**route-policy** *policyname*]

source/mask、*protocol*、**route-policy** 为静态组播路由的关键要素，三个关键要素有一个不同就认为是不同的配置。配置时，首先查询静态组播路由树中是否已有此配置，若有，就修改相应的字段（配置顺序保持不变）；若没有，则添加。每一个网段允许配 8 条不同的配置。

📖 **说明：**

配置了 **ip rpf-route-static** 之后，该组播静态路由并不一定会生效，因为有可能出现无法迭代出接口，或者指定接口 down 等情况。因此，进行此项配置后，建议使用命令 **display multicast routing-table static config** 查看该路由是否配置成功，使用命令 **display multicast routing-table static** 查看该路由是否生效。

8.2.2 配置组播 RPF 路由选择策略

组播 RPF 路由选择策略分为最长匹配规则和优先级规则两种：最长匹配规则是指在“MBGP、组播静态路由、单播路由”三种路由中选择掩码最长的；优先级规则是指在上述三种路由中选择优先级最高（数值最小）。

请在系统视图下进行下列配置。

操作	命令
配置使用最长匹配规则	**ip rpf-longest-match**
恢复缺省配置	**undo ip rpf-longest-match**

缺省情况下，使用优先级规则进行选路。

8.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置后组播的运行情况，验证配置的效果。

操作	命令
显示配置的组播路由	**display multicast routing-table** [*source-address* **mask** *mask-value*] | [*group*-address **mask** *mask-value*] | [**incoming-interface** *interface-name*] *

8.4 配置实例

1. 组网需求

在图 8-2所示的环境中，路由器 Router A 和 Router B 工作在 OSPF 域中，其中路由器 Router A 支持组播，路由器 Router B 上运行单播路由协议，Router C 为域外运行组播路由协议的路由器。要想使 Router A 通过 Router C 获得域外的组播源信息，通过 Router B 获得域内的组播源信息，可以进行下面的配置。

2. 组网图

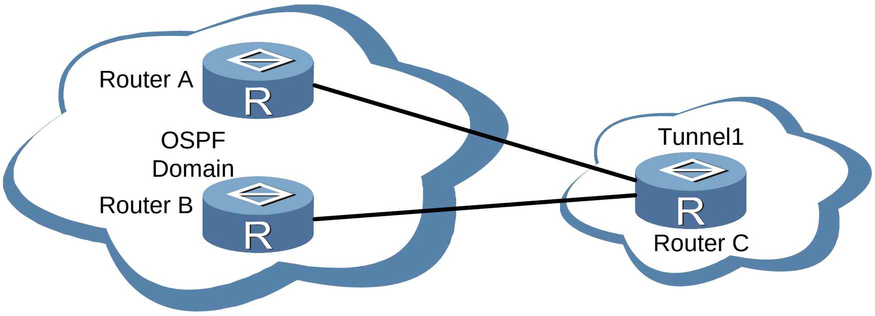


图8-2 组播静态路由典型配置案例

3. 配置步骤

Router A 的配置（仅列出组播静态路由相关的配置）：

```
<Router A> system-view
[Router A] ip rpf-route-static 0.0.0.0 0.0.0.0 ospf null0 preference 255
[Router A] ip rpf-route-static 0.0.0.0 0.0.0.0 tunnel1
```

其中，第一条命令使域内的组播源信息经由 Router B 到达 Router A 时，如果满足匹配规则，则通过 RPF 检查；第二条命令使域外的组播源信息经由 tunnel1 到达 Router A 时通过 RPF 检查。

第9章 配置组播业务

本章包括以下内容：

- 组播业务简介
- 配置内容
- 显示和调试
- 配置实例

9.1 组播业务简介

9.1.1 背景

现有的基于 xDSL（x Digital Subscriber Line）技术的宽带网主要以承载数据业务为主，视频业务还有待发展。制约宽带网用户增长的一个重要因素就是缺乏内容，网络上缺乏象电视一样能够吸引大量用户的内容或运营模式。因此对于电信运营商来说，要想成为真正的全业务运营商，开展具有吸引力的宽带增值业务，吸引更多的用户使用宽带网，解决宽带承载视频业务的问题是当务之急。

MA5200G 支持强大的组播业务，基于组播，可以大规模地开展 IPTV（IP Television）业务。各种端口的无阻塞、低时延（平均时延 20us～80us）线速转发，可以保证视频业务的图像质量。MA5200G 可以对机顶盒 STB（Set Top Box）提供接入认证、安全控制、防攻击等管理能力，提供高带宽的 VoD（Video on Demand）业务点播能力。

9.1.2 原理

华为公司宽带组播视频解决方案的网络模型结构如图 9-1所示：

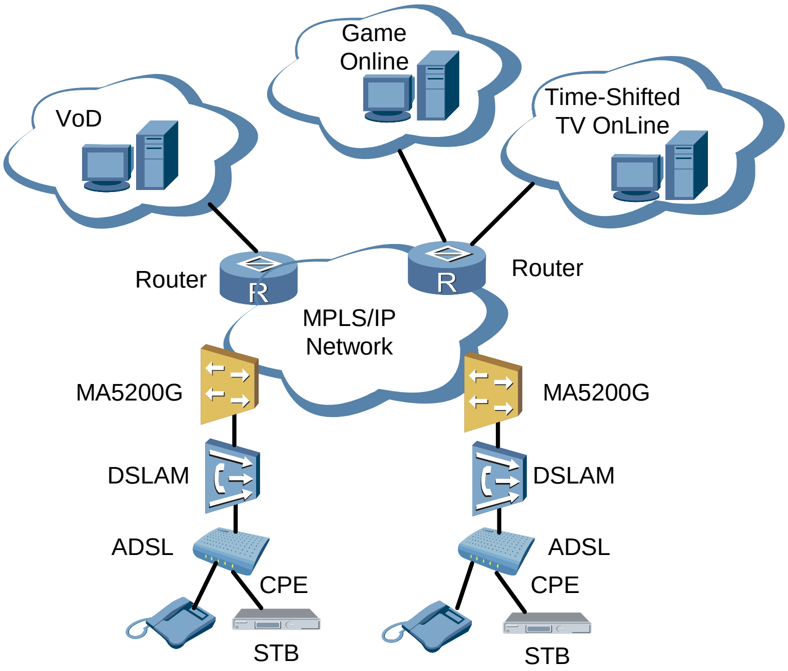


图9-1 组播业务网络模型

组播视频网络由视频业务系统、IP 城域网和家庭网络三个部分组成。

1. 视频业务系统

视频业务系统实现视频业务管理、视频用户管理、视频编码、内容保护（CA/DRM）等功能，并将视频业务流送入城域网。每一个 TV 和广播频道采用 MPEG 编码成一路码流封装到 UDP/IP 报文中，使用一个唯一的 D 类多播 IP 地址。

2. IP 城域网

通过 IP 组播功能将视频业务流送到 IP 城域网，也可以建设独立的视频内容分发网 VDN 将视频内容送到 IP 城域网。VDN 采用基于 GSR（Gigabit Switching Router）和 BRAS（Broadband Remote Access Server）建设。

IP 城域网实现用户宽带上网接入认证管理、视频组加入离开控制功能，并将用户需要的视频流发送给用户。IP 城域网包含 BRAS、组播路由器（Multicast Router）和 AN 设备（包括 DSLAM 和综合接入、L2 Switch 等设备）。

视频流从城域网发送到组播路由器，再送到二层交换机，然后送到 AN 设备（也可以通过专用的视频分发网 VDN 将视频流直接推送到 AN），AN 根据用户的 IGMP 控制报文将视频流发送给需要的用户。

组播路由器实现组播路由协议（如 PIM 协议）和 IGMP。AN 实现 IGMP Proxy 或 IGMP 查询器功能。组播以外的业务，如用户上网业务流、CA/DRM 控制流和单播视频流（VOD）等通过 BRAS 接入城域网。

3. 家庭网络

在用户侧，由于存在 Modem、PC、STB、IAD 等多种设备，这些设备内部互联形成一个家庭局域网。来自 xDSL 线路的视频业务通过 CPE（Customer Premises Equipmen）和家庭网络传送到机顶盒 STB，STB 实现解码功能，解出视频信号送到电视机上显示。

9.2 配置内容

配置组播业务包括以下内容：

- 配置组播组
- 创建组播组列表
- 设置组播组列表引用的组播组
- 创建组播模板
- 设置组播模板引用的组播组列表
- 设置组播模板最多引用的组播组列表数（可选）
- 设置对组播用户进行认证（可选）
- 打开可控组播开关

9.2.1 配置组播组

本任务用于配置组播组的名称、IP 地址、是否预加入组播组。

所谓预加入是指当 MA5200G 下没有该组播组用户时，MA5200G 还保持和组播源的连接，并接收组播报文。缺省情况下，MA5200G 不预加入组播组。

请在 AAA 视图下进行下列配置。

操作	命令
配置组播组	**multicast-group** *group-name* *multicast-group-address* [**prejoin**]
删除组播组	**undo multicast-group** *group-name*

缺省情况下，MA5200G 中未配置组播组。

9.2.2 创建组播组列表

本任务用于创建组播组列表并进入组播组列表视图，如果该组播组列表已经存在，则直接进入组播组列表视图。

请在 AAA 视图下进行下列配置。

操作	命令
创建组播组列表并进入组播组列表视图	**multicast-group-list** *list-name*
删除组播组列表	**undo multicast-group-list** *list-name*

缺省情况下，MA5200G 中未配置组播组列表。

9.2.3 设置组播组列表引用的组播组

本任务用于设置组播组列表所引用的组播组，可按组播组名字或索引进行设置。

请在组播组列表视图下进行下列配置。

操作	命令
设置组播组列表所引用的组播组	**multicast-group** { **name** *group-name* | **start-mg-index** *start-index* [**end-mg-index** *end-index*] }
取消组播组列表所引用的组播组	**undo multicast-group** { **name** *group-name* **|** **start-mg-index** *start-index* [**end-mg-index** *end-index*] }

缺省情况下，组播组列表未引用组播组。

9.2.4 创建组播模板

本任务用于创建组播模板并进入组播模板视图，如果该组播模板已经存在，则直接进入组播模板视图。

请在 AAA 视图下进行下列配置。

操作	命令
创建组播模板并进入组播模板视图	**multicast-profile** *profile-name*
删除组播模板	**undo multicast-profile** *profile-name*

- 当组播模板被域或 BAS 接口引用时，不能被删除。
- 组播模板可引用多个组播组列表。当 BAS 接口或域引用了组播模板，BAS 接口或域的用户将拥有组播模板所引用的组播组列表，进而可加入组播组列表所引用的组播组。
- 缺省情况下，MA5200G 中未配置组播模板。

9.2.5 设置组播模板引用的组播组列表

本任务用于设置组播模板所引用的组播组列表，可按组播组列表名字或索引进行设置。一个组播模板可引用多个组播组列表。

请在组播模板视图下进行下列配置。

操作	命令
设置组播模板所引用的组播组列表	**multicast-group-list** { **name** *list-name* | **start-mglist-index** *start-index* [**end-mglist-index** *end-index*] }
取消组播模板所引用的组播组列表	**undo multicast-group-list** { **name** *list-name* | **start-mglistlist-index** *start-index* [**end-mglist-index** *end-index*] }

缺省情况下，组播模板未引用组播组列表。

9.2.6 设置组播模板最多引用的组播组列表数

本任务用于设置组播模板最多可引用的组播组列表数。

请在组播模板视图下进行下列配置。

操作	命令
设置组播模板最多可引用的组播组列表数	**max-mgroup-num** *number*
恢复可引用的组播组列表数缺省值	**undo max-mgroup-num**

缺省情况下，一个组播模板可引用一个组播组列表。

9.2.7 设置对组播用户进行认证

本任务用于设置对组播用户进行认证。当组播模板设置了对组播用户进行认证后，则所有引用该组播模板的 BAS 接口或域下的用户使用组播业务时，均需要认证。

请在组播模板视图下进行下列配置。

操作	命令
设置对组播用户进行认证	**authentication**
取消对组播用户进行认证	**undo authentication**

缺省情况下， MA5200G 对组播用户进行认证。

9.2.8 打开可控组播开关

本任务用于打开接口的可控组播开关。打开可控组播开关后，接口会依据 IPTV 视图下的配置，对用户加入组播组的权限进行鉴权控制。

请在接口视图下进行下列配置。

操作	命令
打开接口的可控组播开关	**multicast authorization-enable**
关闭接口的可控组播开关	**undo multicast authorization-enable**

缺省情况下，可控组播开关为关闭状态。

9.3 显示和调试

在完成上述配置后，可在任意视图下执行 **display** 命令显示配置组播业务后的运行情况，验证配置的效果。

操作	命令
显示组播组信息	**display multicast-group** [*group-name*]
显示组播组列表信息	**display multicast-group-list** [*list-name*]
显示组播模板信息	**display multicast-profile** [*profile-name*]

9.4 配置实例

1. 组网需求

在如图 9-2所示的环境中，要求从 MA5200G（名称为 Quidway）的 Ethernet 2/0/1 接入、属于 default0 域的用户可以点播地址为 224.0.1.2 的组播组可以点播的节目。

2. 组网图

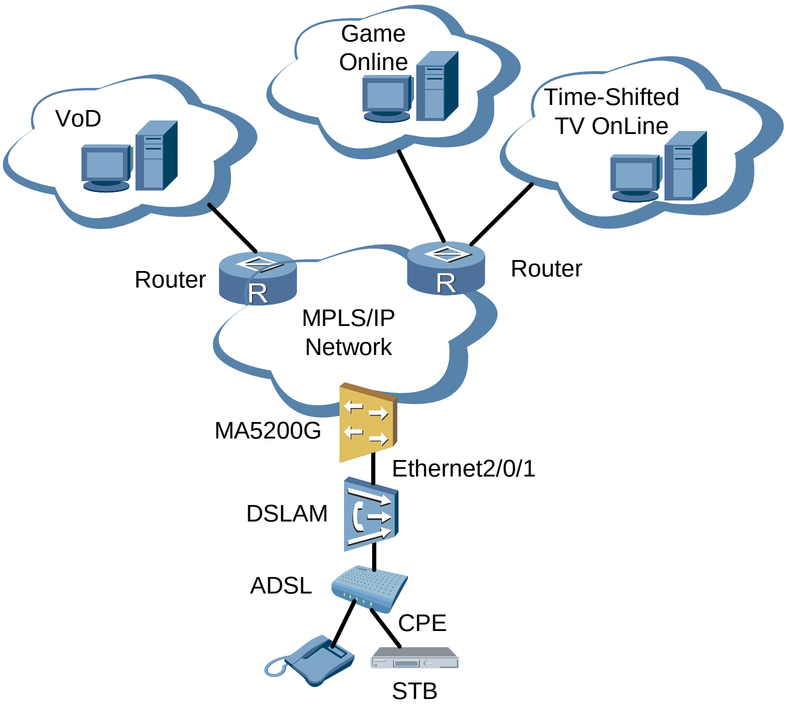


图9-2 组播业务应用配置组网图

3. 配置步骤

配置名为 group1 的组播组。

```
<Quidway> system-view
[Quidway] aaa
[Quidway-aaa] multicast-group group1 224.0.1.2 prejoin
new multicast group created
[Quidway-aaa]
```

创建组播组列表 list1，引用的组播组为 group1。

```
[Quidway-aaa] multicast-group-list list1
[Quidway-aaa-mglist-list1] multicast-group name group1
[Quidway-aaa-mglist-list1] quit
[Quidway-aaa]
```

创建组播模板 profile1，引用的组播组列表为 list1。

```
[Quidway-aaa] multicast-profile profile1
[Quidway-aaa-mprofile-profile1] multicast-group-list name list1
[Quidway-aaa-mprofile-profile1] quit
[Quidway-aaa] quit
[Quidway]
```

打开组播路由开关。

```
[Quidway] multicast routing-enable
```

在用户上线的 default0 域引用组播模板 profile1。

```
[Quidway] aaa
[Quidway-aaa] domain default0
[Quidway-aaa-domain-default0] multicast-profile profile1
[Quidway-aaa-domain-default0] quit
[Quidway-aaa] quit
[Quidway]
```

在以太网接口 2/0/1 下打开可控组播开关。

```
[Quidway] interface Ethernet 2/0/1
[Quidway-Ethernet2/0/1] multicast authorization-enable
```